

RNA-Seq-basierte Isolierung  
des Resistenzgens *Bs4C* aus Paprika

Dissertation der Fakultät für Biologie  
der Ludwig-Maximilians-Universität München

Tina Strauß  
München, 2016

Diese Dissertation wurde angefertigt  
unter der Leitung von Prof. Thomas Lahaye  
am Institut für Genetik  
an der Ludwig-Maximilians-Universität München

Erstgutachter: Prof. Dr. Thomas Lahaye  
Zweitgutachter: Prof. Dr. Martin Parniske

Tag der Abgabe: 25.09.2014

Tag der mündlichen Prüfung: 13.04.2015

## **ERKLÄRUNG**

Hiermit versichere ich ehrenwörtlich, dass meine Dissertation selbständig und ohne unerlaubte Hilfsmittel angefertigt worden ist. Die vorliegende Dissertation wurde weder ganz, noch teilweise bei einer anderen Prüfungskommission vorgelegt.

Ich habe noch zu keinem früheren Zeitpunkt versucht, eine Dissertation einzureichen oder an einer Doktorprüfung teilzunehmen.

München, den 10.04.2016

## Zusammenfassung

Das Paprika Resistenzgen *Bs4C* aus *Capsicum pubescens* vermittelt Resistenz gegenüber *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*)-Stämmen, die den (*transcription activator-like*) TAL-Effektor AvrBs4 exprimieren. Vorangegangene Arbeiten ließen vermuten, dass AvrBs4 die Expression von *Bs4C* transkriptionell induziert. In einem “proof of principle”-Experiment, wurde *Bs4C* unter Verwendung eines RNA-Seq-basierten Ansatzes isoliert. Unter 68 differentiell AvrBs4-induzierten Paprikagenen war jedoch nur eines, das ausschließlich in der resistenten und nicht in der suszeptiblen Akzession induziert war und für das kein Transkript in Abwesenheit von AvrBs4 in der resistenten Akzession nachgewiesen wurde. Kopplungs- und Komplementationsanalysen bestätigten dieses Kandidatengen als das gesuchte Resistenzgen *Bs4C*.

Im *Bs4C*-Promoter konnte ein Effektorbindeelement (EBE) für AvrBs4 identifiziert werden, das notwendig und ausreichend für die AvrBs4-Bindung an und transkriptionelle Aktivierung von *Bs4C* ist. Bindungsstudien ließen erkennen, dass zwei Nukleotidpolymorphismen in der korrespondierenden Region der suszeptiblen Akzession eine stark reduzierte Affinität (10fach) gegenüber AvrBs4 bedingen. Außerdem zeigten GUS-Studien, dass der Promoter des suszeptiblen Allels nicht durch AvrBs4 induzierbar ist. Folglich bestimmt ein Substitutionspolymorphismus von zwei Basenpaaren in den Promotoren des resistenten und suszeptiblen *Bs4C*-Allels über Resistenz oder Suszeptibilität gegenüber AvrBs4-exprimierenden Xanthomonaden.

*Bs4C* kodiert für ein 164-AS großes Protein, das keine Homologie zu Proteinen mit bekannter Funktion aufweist. *In silico* Proteinstrukturanalysen sagen vier Transmembranhelices in *Bs4C* vorher und demzufolge stellt es einen neuen Typ von Exekutorproteinen dar, welcher Resistenz gegen TALE-exprimierende *Xanthomonas*-Stämme vermittelt.

Zudem konnten Sequenzanalysen mindestens ein Homolog in *C. pubescens* und mindestens sieben Homologe in *C. annuum* identifizieren. Interessanterweise kodieren die meisten von ihnen aufgrund von Nukleotidaustauschen, Leserahmenverschiebungen und Insertions/Deletionspolymorphismen nicht für Vollängen *Bs4C*-ähnliche Proteine. Folglich könnten diese homologen Sequenzen duplizierte Gene repräsentieren, die nicht mehr funktional sind.

## Summary

The *Capsicum pubescens* gene *Bs4C* confers resistance to *Xanthomonas campestris* pv. *vesicatoria* (*Xcv*) strains expressing the transcription activator-like (TAL)-effector AvrBs4. Preliminary studies indicated that AvrBs4 transcriptionally activate the expression of *Bs4C*. In a poof-of-principle experiment we isolated *Bs4C* using a RNA-Seq-based approach. Among 68 pepper genes that were transcriptionally activated by AvrBs4 only one candidate gene was exclusively induced in the resistant and not in the susceptible as well as no transcript level was detectable in the absence of AvrBs4 in the resistant pepper accession. Linkage mapping and complementation assays confirmed that this candidate gene was the searched-for resistance gene *Bs4C*.

Within the *Bs4C* promoter an effector binding element (EBE) for AvrBs4 was identified, that is required and sufficient for AvrBs4-binding to and transcriptional activation of *Bs4C*. Analysis to the corresponding region in a susceptible pepper accession revealed a two nucleotide polymorphism resulting in a lower affinity (10 fold) of AvrBs4 to the promoter region of the susceptible allele. Furthermore, transient GUS assays showed that the promoter of the susceptible allele is not AvrBs4-inducible. Thus, a 2-bp substitution-polymorphism in the promoters of the resistant and susceptible allele of *Bs4C* determines resistance or susceptibility to xanthomonads containing AvrBs4.

*Bs4C* encodes for a 164-amino acid protein showing no homology to proteins with known function. *In silico* protein predictions revealed four potential transmembrane helices in *Bs4C* representing a novel type of executor protein that mediates resistance to a TALE-expressing *Xanthomonas* strain.

Sequence analyses uncovered at least one homolog in *C. pubescens* as well as at least seven homologs in *C. annuum*. Interestingly, most of them do not encode for full-length *Bs4C*-like proteins due to nucleotide exchanges, frame shift or insertion-/deletion-polymorphisms. Thus, these homologous sequences could represent duplicated genes that lost their function.



## Übersicht über Publikationen

**Strauß T, Van Poecke R, Strauß A, Römer P, Minsavage GV, Singh S, Wolf C, Strauß A, Kim S, Lee H-A, Yeom S-I, Parniske M, Stall RE, Jones JB, Choi D, Prins M, Lahaye T.**

RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large crop genome. **Proc Natl Acad Sci.** 2012 Oct; 109: 19480 - 19485

**Römer P, Recht S, Strauss T, Elsaesser J, Schornack S, Boch J, Wang S, Lahaye T.**

Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. **New Phytol.** 2010 Sep; 187(4): 1048-57

**Römer P, Strauss T, Hahn S, Scholze H, Morbitzer R, Grau J, Bonas U, Lahaye T.**

Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper *Bs3* alleles. **Plant Physiol.** 2009 Aug; 150(4): 1697-712

**Römer P, Hahn S, Jordan T, Strauss T, Bonas U, Lahaye T.**

Plant pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. **Science.** 2007 Oct 26; 318(5850): 645-8

# Inhaltsverzeichnis

Zusammenfassung .....	I
Summary .....	II
Übersicht über Publikationen .....	III
Abbildungsverzeichnis .....	VII
Tabellenverzeichnis .....	IX
Abkürzungsverzeichnis .....	X
<b>1 Einleitung .....</b>	<b>1</b>
1.1 Pflanzliche Abwehrmechanismen.....	1
1.2 Resistenzproteine gegen bakterielle Effektoren.....	2
1.3 Xanthomonaden - ökonomisch bedeutende Bakterien .....	3
1.4 <i>Xanthomonas</i> TAL-Effektoren - Struktur und TALE-Code .....	4
1.5 <i>Xanthomonas</i> TAL-Effektoren - Virulenzfunktion .....	7
1.6 TAL-Effektor R-Proteine - Induktion pflanzlicher Abwehrreaktionen .....	9
1.7 Isolierung von R-Genen mittels <i>map-based cloning approach</i> (kartengestützten Ansatz) .....	12
1.8 Arbeitshypothese: Isolierung eines TAL-induzierten R-Gens mittels RNA-Seq-basierter Transkriptomanalysen .....	12
1.9 Zielstellung der Arbeit .....	14
<b>2 Material und Methoden .....</b>	<b>15</b>
2.1 Chemikalien, Geräte und Verbrauchsmaterialien .....	15
2.2 Verwendete Kits.....	15
2.3 Verwendete Nährmedien und deren Zusammensetzung.....	15
2.4 Verwendete Lösungen und Puffer .....	16
2.5 Verwendete Größenstandards .....	16
2.6 Verwendete Antibiotika (Stammkonzentration und Verdünnungen) .....	17
2.7 Verwendete Vektoren .....	17
2.8 Zusammenstellung verwendeter Bakterienstämme .....	17
2.9 Zusammenstellung verwendeter Oligonukleotide .....	18
2.10 Pflanzenmaterial .....	18
2.11 Methoden .....	19
2.11.1 Bakterien .....	19
2.11.1.1 Kultivierung von Bakterien .....	19
2.11.1.2 Transformation chemisch-kompetenter Bakterien .....	19
2.11.1.3 Transfer von Plasmid-DNA in elektrokompetente Zellen.....	20
2.11.1.4 Dauerhafte Lagerung von Bakterien.....	20
2.11.1.5 Isolation von Plasmid-DNA aus Bakterien.....	20
2.11.1.6 Transiente Expression in <i>N. benthamiana</i> durch <i>A. tumefaciens</i> .....	20
2.11.1.7 GUS-Analysen nach transienter Expression in <i>N. benthamiana</i> .....	21
2.11.1.8 Inokulationsexperimente mit <i>Xcv</i> .....	21
2.12 DNA-Methoden .....	22
2.12.1 Agarose-Gelelektrophorese.....	22
2.12.2 Polymerasekettenreaktion (PCR).....	22
2.12.2.1 Kolonie-PCR.....	23
2.12.3 Restriktionsanalysen .....	23
2.12.4 GATEWAY-basierte Klonierung von DNA-Fragmenten .....	24
2.12.5 EMSA- und MST-Analysen .....	24
2.12.6 Isolierung von Gesamt-RNA .....	24
2.12.6.1 Qiagen RNeasy Plant Mini Kit (für RT-PCR).....	24
2.12.6.2 Trizol-Extraktion .....	25
2.12.7 Isolierung genomischer Pflanzen-DNA .....	25
2.12.8 Konzentrationsbestimmung der Nukleinsäuren .....	26
2.12.9 cDNA Synthese für RT-PCR-Analysen.....	26
2.12.10 RACE ( <i>rapid amplification of cDNA ends</i> ) .....	27
2.12.11 Sequenzierung.....	27
2.12.12 Restriktionsanalysen und Ligation.....	27

2.12.13	Cut-Ligation .....	27
2.12.14	Erstellung der designer TALEs (dTALEs) .....	28
2.13	<i>C. annuum</i> cv. CM334-Genomprojekt.....	28
2.14	Erstellung einer <i>C. pubescens</i> cDNA-Bibliothek (Invitrogen).....	29
2.14.1	Organisation der cDNA-Bibliothek .....	29
2.15	Illumina-basierte Transkriptanalysen.....	30
2.16	Verwendete Softwareprogramme .....	31
<b>3</b>	<b>Ergebnisse .....</b>	<b>33</b>
3.1	Isolierung AvrBs4-induzierter Genen mittels Transkriptomprofiling .....	35
3.2	RT-PCR-Analysen zur Bestätigung des Kandidatengens 12600 als AvrBs4-induziertes Gen .....	36
3.3	Die genetische Kartierung des Kandidaten 12600 bestätigt die Kopplung mit der AvrBs4-induzierten Resistenzreaktion .....	39
3.4	Das <i>Bs4C</i> -Kandidatengens 12600 vermittelt die Erkennung von AvrBs4.....	40
3.4.1	Isolierung und vergleichende Sequenzierung des <i>Bs4C</i> -Kandidatengens aus der suszeptiblen und resistenten <i>C. pubescens</i> -Akzession.....	40
3.4.2	Komplementationanalysen bestätigen den Kandidaten 12600 als Resistenzgen <i>Bs4C-R</i> .....	41
3.5	Die Überexpression der <i>Bs4C</i> -Allele resultiert in einer HR.....	41
3.6	Das <i>EBE<sub>AvrBs4</sub>Bs4C-R</i> vermittelt die AvrBs4-Erkennung .....	43
3.7	AvrBs4 bindet mit höherer Affinität an das <i>EBE<sub>AvrBs4</sub>Bs4C-R</i> im Vergleich zum <i>EBE<sub>AvrBs4</sub>Bs4C-S</i> .....	45
3.8	Analyse weiterer AvrBs4-induzierter Kandidaten.....	47
3.8.1	Kandidat 630 .....	49
3.8.2	Kandidat 15502 .....	51
3.8.3	Kandidat 8248 .....	51
3.9	Assemblierung eines designer TALE zur Induktion von <i>Bs4C-S</i> .....	52
3.10	Analyse von <i>Bs4C</i> -Homologen in <i>Capsicum annuum</i> .....	54
3.10.1	Identifizierung von <i>Bs4C</i> -Homologen im assemblierten Genom von <i>C. annuum</i> cv. CM334 .....	54
3.11	Analyse der <i>Bs4C</i> -Homologen <i>CaBs4C.1</i> bis <i>CaBs4C.7</i> .....	55
3.11.1	ORF und ORF-korrespondierende Sequenzen.....	55
3.11.2	EBE in den Promotersequenzen von <i>CaBs4C.1</i> - <i>CaBs4C.5</i> .....	56
3.11.3	Funktionale Analyse der <i>Bs4C</i> -Homologen aus <i>C. annuum</i> .....	57
3.12	Analyse von <i>Bs4C</i> -Paralogen in <i>Capsicum pubescens</i> .....	59
3.12.1	Identifikation von <i>Bs4C</i> -Paralogen in <i>C. pubescens</i> .....	59
3.12.2	Sequenzanalyse von <i>Bs4C</i> -Paralogen in <i>C. pubescens</i> .....	59
3.12.3	Funktionale Analyse von <i>CpBs4C.2-R</i> und <i>CpBs4C.2-S</i> .....	60
3.12.4	RT-PCR-Analysen für <i>CpBs4C.2-R</i> und <i>CpBs4C.2-S</i> .....	60
3.13	Identifikation von <i>Bs4C-R</i> -Homologen in Solanaceae-Spezies .....	62
3.13.1	Phänotypische Analyse .....	62
3.13.2	Sequenzanalyse und Vergleich der EBE.....	62
3.13.3	Sequenzanalyse und Vergleich der <i>Bs4C-R</i> -homologen Sequenzen .....	64
<b>4</b>	<b>Diskussion.....</b>	<b>65</b>
4.1	RNA-Seq kann zur Isolierung von TALE-induzierten Resistenzgenen verwendet werden.....	65
4.2	Kritische Parameter für RNA-Seq-basierte Isolierung von TAL-induzierten <i>R</i> -Genen .....	67
4.2.1	Auswahl des Zeitpunktes der Probennahme .....	67
4.2.2	Erstellung der cDNA-Bibliothek .....	67
4.2.3	Nutzung von Cycloheximid zur Identifizierung des <i>R</i> -Gens als direktes Zielgen unter der Gesamtheit der TAL-induzierten Gene.....	69
4.2.4	Vorhersageprogramme allein genügen nicht zur Isolierung von TALE-induzierten Resistenzgenen.....	71

4.3	<i>Bs4C-R</i> gehört zur Gruppe der TALE-induzierten Resistenzgene und kodiert für ein putatives Transmembranprotein.....	73
4.4	Potentielle AvrBs4-Suszeptibilitätsgene in <i>C. pubescens</i> .....	76
4.5	<i>Bs4C</i> -Homologe in <i>Capsicum</i> - Funktion außerhalb der Resistenz?.....	78
<b>5</b>	<b>Literatur .....</b>	<b>81</b>
<b>6</b>	<b>Anhang.....</b>	<b>90</b>
	Danksagung .....	132

## Abbildungsverzeichnis

Abb. 1 Krankheitssymptome verursacht durch verschiedene <i>Xanthomonas</i> -Stämme..	4
Abb. 2 Struktur von TAL-Effektoren, RVDs und TALE-Code.....	6
Abb. 3 Virulenz- und Resistenzinduktion durch TAL-Effektoren in der Pflanze .....	9
Abb. 4 Übersicht über TAL- <i>R</i> -Gen-Interaktionen .....	11
Abb. 5 Die Erkennung von AvrBs4 in <i>C. pubescens</i> PI 235047 durch <i>Bs4C</i> ist abhängig von der Aktivierungsdomäne.....	14
Abb. 6 Schema der Organisation der cDNA-Bibliothek zur Analyse mittels PCR ....	30
Abb. 7 Bestätigung des <i>Bs4C-R</i> -Kandidaten 12600 als direktes AvrBs4-Zielgen mittels sq RT-PCR-Analysen.....	38
Abb. 8 Transkriptnachweis des <i>Bs4C</i> -Kandidatengens 15 Stunden nach AvrBs4 Induktion .....	38
Abb. 9 Der <i>Bs4C-R</i> -Kandidat ist genetisch gekoppelt mit der AvrBs4-induzierten Immunantwort .....	39
Abb. 10 Genstruktur von <i>Bs4C-R</i> vergleichend zu <i>Bs4C-S</i> .....	40
Abb. 11 Das <i>Bs4C-R</i> -Kandidatengens vermittelt die Erkennung von AvrBs4 .....	41
Abb. 12 Die konstitutive Expression der <i>Bs4C</i> -Allele resultiert in der Ausbildung einer HR .....	42
Abb. 13 <i>Bs4C-R</i> enthält ein Effektorbindeelement (EBE), welches notwendig und ausreichend für die transkriptionelle Aktivierung durch AvrBs4 ist .....	43
Abb. 14 Die Nukleotidsequenz des <i>EBE<sub>AvrBs4</sub>Bs4C-R</i> weist fünf Polymorphismen zu dem anhand des TALE-Codes vorhergesagten <i>EBE<sub>AvrBs4</sub>*</i> auf .....	44
Abb. 15 EMSA und EMSA-Kompetitionsstudien sowie MST-Analysen zeigen, dass AvrBs4 eine höhere Affinität zum <i>EBE<sub>AvrBs4</sub>Bs4C-R</i> im Vergleich zum <i>EBE<sub>AvrBs4</sub>Bs4C-S</i> aufweist.....	47
Abb. 16 Auswahl einiger sq RT-PCR-Analysen in An- und Abwesenheit von CHX zur Bestimmung von AvrBs4-direkt induzierten Genen .....	49
Abb. 17 Vergleich und funktionaler Test der EBE AvrBs4-induzierter Gene.....	50

Abb. 18 Der dTALE <i>EBE Bs4C-S</i> aktiviert spezifisch den <i>Bs4C-S</i> -Promoter aber nicht den <i>Bs4C-R</i> -Promoter.....	53
Abb. 19 Anordnung der <i>Bs4C-R</i> -Homologen auf Chromosom 1 des <i>C. annuum</i> cv. CM334-Genoms .....	54
Abb. 20 Das <i>EBE<sub>AvrBs4</sub>CaBs4C.1</i> wird nicht durch AvrBs4 induziert.....	57
Abb. 21 Die konstitutive Expression des <i>C. annuum Bs4C</i> -Ortholog <i>CaBs4C.1</i> resultiert in einer HR .....	58
Abb. 22 Sequenzidentitäten verschiedener <i>Bs4C-R</i> -Homologen zueinander.....	60
Abb. 23 Die konstitutive Expression der <i>C. pubescens Bs4C</i> -Paraloge <i>CpBs4C.2-R</i> und <i>CpBs4C.2-S</i> resultiert in einer HR .....	60
Abb. 24 Vergleich der EBE von <i>CpBs4C.2-R</i> mit denen von <i>Bs4C-S</i> und den <i>C. annuum</i> -Homologen <i>CaBs4C.1</i> und <i>CaBs4C.2</i> .....	61
Abb. 25 Die <i>Bs4C</i> -Paraloge <i>CpBs4C.2-R</i> und <i>CpBs4C.2-S</i> sind nicht durch AvrBs4 induziert.....	62
Abb. 26 Unterscheidung zwischen direkten (A) und indirekten (B) AvrBs4-Zielgenen durch Cycloheximid .....	69
Abb. 27 Proteinvorhersagen für <i>Bs4C-R</i> und Xa10 mit dem Programm TMHMM...	74
Abb. 28 Vergleich der EBE der <i>Bs4C</i> -Allele aus <i>C. pubescens</i> und <i>C. annuum</i> .....	80

## Tabellenverzeichnis

Tab. 1 Das Expressionsmuster des <i>Bs4C</i> -Kandidatengen 12600.....	36
Tab. 2 Identifikation von <i>Bs4C-R</i> -Homologen in <i>C. annuum</i> cv. CM334 und deren Vergleich mit <i>Bs4C-R</i> und <i>Bs4C-S</i> .....	56
Tab. 3 Identifikation <sup>1)</sup> und Vergleich potentieller EBE der <i>Bs4C-R</i> -Homologen aus <i>C. annuum</i> cv. CM334 .....	56
Tab. 4 Übersicht der EBEs in den analysierten <i>Capsicum</i> -Spezies .....	64
Tab. 5 Vergleich der Daten der RNA-Seq- und sq RT-PCR-Analysen .....	72

## Abkürzungsverzeichnis

A	- Adenin
Abb.	- Abbildung
AD	- <i>acidic activation domain</i> (saure Aktivierungsdomäne)
<i>A. tumefaciens</i>	- <i>Agrobacterium tumefaciens</i>
AS	- Aminosäure(n)
ATG	- Adenin-Thymin-Guanin-Triplett, Startkodon
<i>avr</i> , Avr	- <i>avirulence</i> (Avirulenz)
BAC	- <i>bacterial artificial chromosome</i>
Bats	- <i>Burkholderia</i> TAL-like
bp	- <i>base pairs</i> (Basenpaar(e))
<i>Bs</i>	- <i>bacterial spot</i>
bspw.	- beispielsweise
BSR	- <i>base specifying residue</i>
bzgl.	- bezüglich
C	- Cytosin
<i>C.</i>	- <i>Capsicum</i>
<i>C. annuum</i>	- <i>Capsicum annuum</i>
CAPS	- <i>cleaved amplified polymorphic sequence</i>
CC	- <i>coiled coil</i>
cDNA	- <i>complementary DNA</i>
CDS/KDS	- <i>coding sequence</i> (kodierende Sequenz)
cfu	- <i>colony-forming units</i> (Kolonie-bildende Einheiten)
CHX	- <i>cycloheximide</i> (Cycloheximid)
<i>CMV</i>	- <i>Cauliflower Mosaic Virus</i> (Blumenkohlmosaik-Virus)
<i>C. pubescens</i>	- <i>Capsicum pubescens</i>
C-Terminus	- carboxyterminal
cv.	- cultivar (Kultivar)
DNA	- <i>deoxyribonucleic acid</i> (Desoxyribonukleinsäure)
dpi	- <i>days post inoculation</i> (Tage nach Inokulation)
dTALE	- <i>designer TALE</i>
EBE	- Effektorbindeelement(e) ( <i>effector binding element</i> )
<i>E. coli</i>	- <i>Escherichia coli</i>
ECW	- Early Calwonder
EF1 $\alpha$	- <i>elongation factor 1 <math>\alpha</math></i> (Elongations-Faktor 1 $\alpha$ )
EMSA	- <i>electrophoretic mobility shift assay</i>
ETI	- <i>effector triggered immunity</i>
EtOH	- Ethanol
FGP	- Forschungsgruppenpraktikum
G	- Guanin
gDNA	- genomic DNA (genomische DNA)
GFP	- <i>green fluorescent protein</i> (grün fluoreszierendes Protein)
GUS	- $\beta$ -Glucuronidase



<i>G. max</i>	- <i>Glycine max</i>
His	- Histidin
hpi	- <i>hours post inoculation</i> (Stunden nach Infiltration)
HR	- <i>hypersensitive response</i> (hypersensitive Reaktion)
kb	- <i>kilobase</i> (Kilobasen)
KDS	- kodierende Sequenz
LPS	- Lipopolysaccharide
LRR	- <i>leucine rich repeat</i>
<i>N. benthamiana</i>	- <i>Nicotiana benthamiana</i>
n.g.	- nicht getestet
NLS	- <i>nuclear localization signal</i> (Kernlokalisationssignale)
nt	- nucleotides (Nukleotid(e))
N-Terminus	- aminoterminal
M	- Molar
MAPK	- <i>mitogen-associated protein kinase</i>
MBP	- Methyl-CpG-bindende Proteinen
min	- Minute(n)
ml	- Milliliter
µl	- Microliter
mM	- millimolar
MST	- <i>microscale thermophoresis</i>
NBS	- <i>nucleotide binding site</i>
nM	- nanomolar
ORF	- <i>open reading frame</i> (offener Leserahmen)
<i>O. sativa</i>	- <i>Oryzae sativa</i>
PAGE	- <i>polyacrylamide gel electrophoresis</i> (Polyacrylamid-Gelelektrophorese)
PAMP	- <i>pathogen associated molecular patterns</i>
PCR	- <i>polymerase chain reaction</i> (Polymerase-Kettenreaktion)
PRR	- <i>plant recognition receptor</i>
PTI	- <i>PAMP-triggered immunity</i>
RACE	- <i>rapid amplification of cDNA ends</i> (schnelle Amplifikation von cDNA-Enden)
RFLP	- <i>restriction fragment length polymorphism</i> (Restriktionsfragmentlängenpolymorphismus)
R-Gen	- <i>resistance gene</i> (Resistenzgen)
Rip	- <i>Ralstonia injected protein</i>
rpm	- <i>rounds per minute</i>
RT	- Raumtemperatur
RT-PCR	- <i>reverse transcription</i> PCR (Reverse Transkriptase-PCR)
RVD	- <i>repeat variable diresidues</i>
sq	- semiquantitativ(e)
<i>S. tuberosum</i>	- <i>Solanum tuberosum</i>
<i>S. lycopersicum</i>	- <i>Solanum lycopersicum</i>

T	- Thymin
Tab.	- Tabelle
TALE	- <i>transcription activator-like effector</i>
T-DNA	- <i>transfer DNA</i> (transferierte DNA)
TIR	- <i>Toll/Interleukin-1 receptor</i>
<i>uidA</i>	- GUS-Gen ( $\beta$ -Glucuronidase)
UTR	- <i>untranslated region</i>
WT	- Wildtyp
<i>Xac</i>	- <i>Xanthomonas axonopodis</i> pv. <i>citri</i>
<i>Xcc</i>	- <i>Xanthomonas campestris</i> pv. <i>campestris</i>
<i>Xcv</i>	- <i>Xanthomonas campestris</i> pv. <i>vesicatoria</i>
<i>Xoc</i>	- <i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>
<i>Xoo</i>	- <i>Xanthomonas oryzae</i> pv. <i>oryzae</i>
X-Gluc	- Cyclohexylammoniumsalz der 5-Brom-4-chlor-3-indolyl- $\beta$ -D-glucuronsäure

# 1 Einleitung

## 1.1 Pflanzliche Abwehrmechanismen

Pflanzen sind ständig einer Vielzahl äußerer Einflüssen ausgesetzt. Um sich gegen biotische Stressoren, wie Fraßfeinden und Pathogene zu schützen, haben sie verschiedene Mechanismen entwickelt. Die passive Abwehr erfolgt dabei durch strukturelle oder chemische Barrieren, die das Eindringen in die Pflanze verhindern. Beispiele hierfür sind die Kutikula, Zellwände und Haare auf der Blattoberfläche sowie antimikrobiell wirkende Sekundärmetabolite, die sich in der Zellwand (hydrolytische Enzyme) oder in anderen Bereichen (Phytoanticipine) befinden (Schaller, 2002). Gelingt es dem Pathogen diese ersten Barrieren zu überwinden und in die Pflanze einzudringen, werden weitere Abwehrmechanismen aktiviert. Die erste Abwehrlinie wird vermittelt durch die Erkennung von sogenannten PAMPs (*pathogen associated molecular patterns*) wie beispielsweise flg22 (Flagellin), EF-Tu (Elongationsfaktor), LPS (Lipopolysaccharide) und Peptidoglykan von Bakterien, Xylanase und Chitin von Pilzen sowie Heptaglucoside von Oomyceten durch oberflächenständige Rezeptoren (PRR, *plant recognition receptor*). Diese Form der Abwehr wird auch als PTI (*PAMP-triggered immunity*) bezeichnet (Dodds and Rathjen, 2010; Jones and Dangl, 2006). PRRs bestehen aus einer Transmembrandomäne und einer extrazellulären LRR-Domäne (*leucine rich repeat*). Die Erkennung der PAMPs durch die PRRs kann verschiedene Reaktionen hervorrufen, die einzeln oder in Kombination die Vermehrung und Ausbreitung des Pathogens eingrenzen/verhindern soll (Jones and Dangl, 2006). Die PTI kann mit der Produktion von reaktiven Sauerstoffspezies einhergehen oder auch der Akkumulation von Kallose. Weitere typische Reaktionen nach der Erkennung von PAMPs sind die Aktivierung von MAP (*mitogen-associated protein*)-Kinasen, die Expression von *defense related genes* oder die Schließung der Stomata (Dodds and Rathjen, 2010; Jones and Dangl, 2006). Einigen Pathogenen ist es möglich, sich der Erkennung zu entziehen oder die PTI zu unterdrücken (He et al., 2007; Zhou and Chai, 2008). Pflanzenpathogene Bakterien schleusen mittels Typ-III-Sekretionssystem sogenannte Effektoren in die Pflanzenzelle, von denen einige in der Lage sind die PTI zu unterbinden (Dodds and Rathjen, 2010; He et al., 2007; Jones and Dangl, 2006). Ein Beispiel dafür ist der *Pseudomonas*-Effektor AvrPto, welcher in suszeptiblen

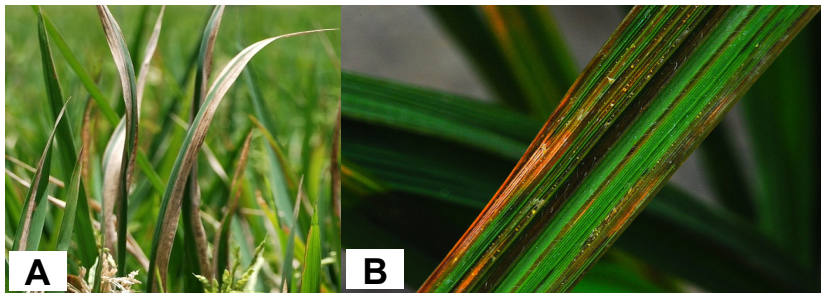
Pflanzen (Tomaten) mit dem PRR FLS2 interagiert und somit die Erkennung von Flagellin und die PTI unterdrückt (Xiang et al., 2008). In Anpassung an bakterielle Effektoren haben Pflanzen im Laufe ihrer Entwicklung eine zweite Abwehrlinie mit spezifischen Immunrezeptoren für diese Effektoren ausgebildet, die als ETI (*effector triggered immunity*) bezeichnet wird. Die Erkennung der Effektoren erfolgt durch Resistenzgenprodukte (*R*-Genprodukte) und resultiert meist in einer lokalen Zelltodreaktion an der Infektionsstelle, in der sogenannten hypersensitiven Reaktion (HR). Die Zelltodreaktion verhindert durch Nahrungsentzug das Wachstum und die weitere Verbreitung des Pathogens (Dodds and Rathjen, 2010; Jones and Dangl, 2006).

## 1.2 Resistenzproteine gegen bakterielle Effektoren

*R*-Gene vermitteln Resistenz gegenüber einer Vielzahl bakterieller Effektoren. Die größte Klasse der *R*-Gene stellen die NBS-LRR Gene dar, welche für zytoplasmatische Proteine kodieren. Sie enthalten eine Nukleotid-Bindestelle (*nucleotide binding site*, NBS) im zentralen Bereich des Polypeptides und einer C-terminalen LRR-Domäne (*leucine rich repeat*) (Dodds and Rathjen, 2010). Basierend auf der Struktur der N-terminalen Domäne können pflanzliche NBS-LRRs entweder in TIR-NBS-LRR- oder CC-NBS-LRR-Proteine differenziert werden (Dodds and Rathjen, 2010). Die TIR-Domäne zeigt Homologie zu der intrazellulären TIR (*Toll/Interleukin-1 receptor*)-Domäne aus Säugern und zu bekannten Vertreter der TIR-NBS-LRRs zählen *RPS4* aus Arabidopsis und *Bs4S* aus Tomate (Gassmann et al., 1999; Schornack et al., 2004). Die CC-NBS-LRRs besitzen eine putative CC (*coiled coil*)-Domäne im N-Terminus und *RPS2* und *RPM1* aus Arabidopsis, welches Resistenz gegenüber *Pseudomonas syringae* vermittelt, können in diese Klasse eingeordnet werden (McHale et al., 2006). Die Erkennung der Effektoren durch NBS-LRR-Proteine kann entweder durch direkte oder indirekte Interaktion erfolgen und löst eine komplexe Abwehrreaktion aus, die meistens mit einer HR einhergeht und somit die Ausbreitung des Pathogens einschränkt (Dodds and Rathjen, 2010). Von der Gruppe der NBS-LRR-Proteine unterscheiden sich deutlich die R-Proteine, die Resistenz gegen TAL (*Transcription Activator-Like*)-Effektoren, einer spezifischen Klasse an Typ-III-Effektorproteinen aus Bakterien der Gattung *Xanthomonas*, vermitteln (Boch and Bonas, 2010).

### 1.3 Xanthomonaden – ökonomisch bedeutende Bakterien

Xanthomonaden gehören zu den Gram-negativen Bakterien und können eine Vielzahl von Pflanzenarten befallen, zu denen auch viele wichtige Kulturpflanzen zählen. Sie sekretieren Exopolysaccharide (Xanthan) und sind durch gelbe Pigmente gekennzeichnet. Die Bakterien werden über Regenwasser und Tautropfen in der Natur verbreitet und gelangen durch natürliche Öffnungen (bspw. Stomata) oder Verwundungen in die Pflanze (Agrios, 1997; Boch and Bonas, 2010). Dabei bevorzugen verschiedene Pathovaren unterschiedliche pflanzliche Gewebe zur Vermehrung. Während sich einige Pathovaren wie *Xanthomonas campestris* pv. *vesicatoria* (Xcv), *Xanthomonas oryzae* pv. *oryzicola* (Xoc) und *Xanthomonas axonopodis* pv. *citri* (Xac) örtlich begrenzt im Interzellularraum (Apoplasten) vermehren, dringen andere, wie *Xanthomonas campestris* pv. *campestris* (Xcc) und *Xanthomonas oryzae* pv. *oryzae* (Xoo) in das Xylem vor und verbreiten sich durch dieses systematisch in der Pflanze. An den Infektionsstellen bilden sich im Krankheitsverlauf wässrige Läsionen (*water-soaked lesions*) und/oder Chlorosen aus, die dann nekrotisch werden. Auch kann es zu einem schnellen Welkprozess kommen, ausgelöst durch ein Xanthan-„verstopftes“ vaskuläres Gewebe (Boch and Bonas, 2010). Xanthomonaden verursachen in Anbaugebieten, in denen sie optimale Bedingungen vorfinden, wie bspw. warm-feuchte Regionen, erhebliche Schäden. Ein eindrucksvolles Beispiel dafür ist *Xanthomonas axonopodis* pv. *citri*, der in Florida (USA) Zitruskrebs (*citrus canker*) hervorruft und aufgrund geschädigter Früchte und daraus resultierender Umsatzeinbußen einen großen ökonomischen Schaden verursacht. Außerdem erfordert er den Einsatz von Pestiziden und die Abholzung von befallenen Pflanzen (Schornack et al., 2013). Ebenso bedingt *Xanthomonas campestris* pv. *vesicatoria*, der Erreger der bakteriellen Fleckenkrankheit (*bacterial spot disease*) auf Paprika (*Capsicum* spp.) und Tomate (*Lycopersicon* spp.) in Florida große Ernteverluste (Jones et al., 1998). Auch Reispflanzen, welche die Nahrungsgrundlage für einen großen Teil der Weltbevölkerung ist, wird von *Xanthomonas*-Stämmen befallen. *Xanthomonas oryzae* pv. *oryzicola* und *Xanthomonas oryzae* pv. *oryzae* bewirken erhebliche Verluste durch die Krankheiten bakterielle Weißblättrigkeit (*bacterial blight*, Abb. 1A) und bakterielle Streifenkrankheit (*bacterial leaf streak*, Abb. 1B) (Nino-Liu et al., 2006).



**Abb. 1 Krankheitssymptome verursacht durch verschiedene *Xanthomonas*-Stämme**

A) Bakterielle Weißblättrigkeit (*bacterial blight*) auf Reis - *Xanthomonas oryzae* pv. *oryzae* (*Xoo*)

B) Bakterielle Streifenkrankheit (*bacterial leaf streak*) auf Reis - *Xanthomonas oryzae* pv. *oryzicola* (*Xoc*)

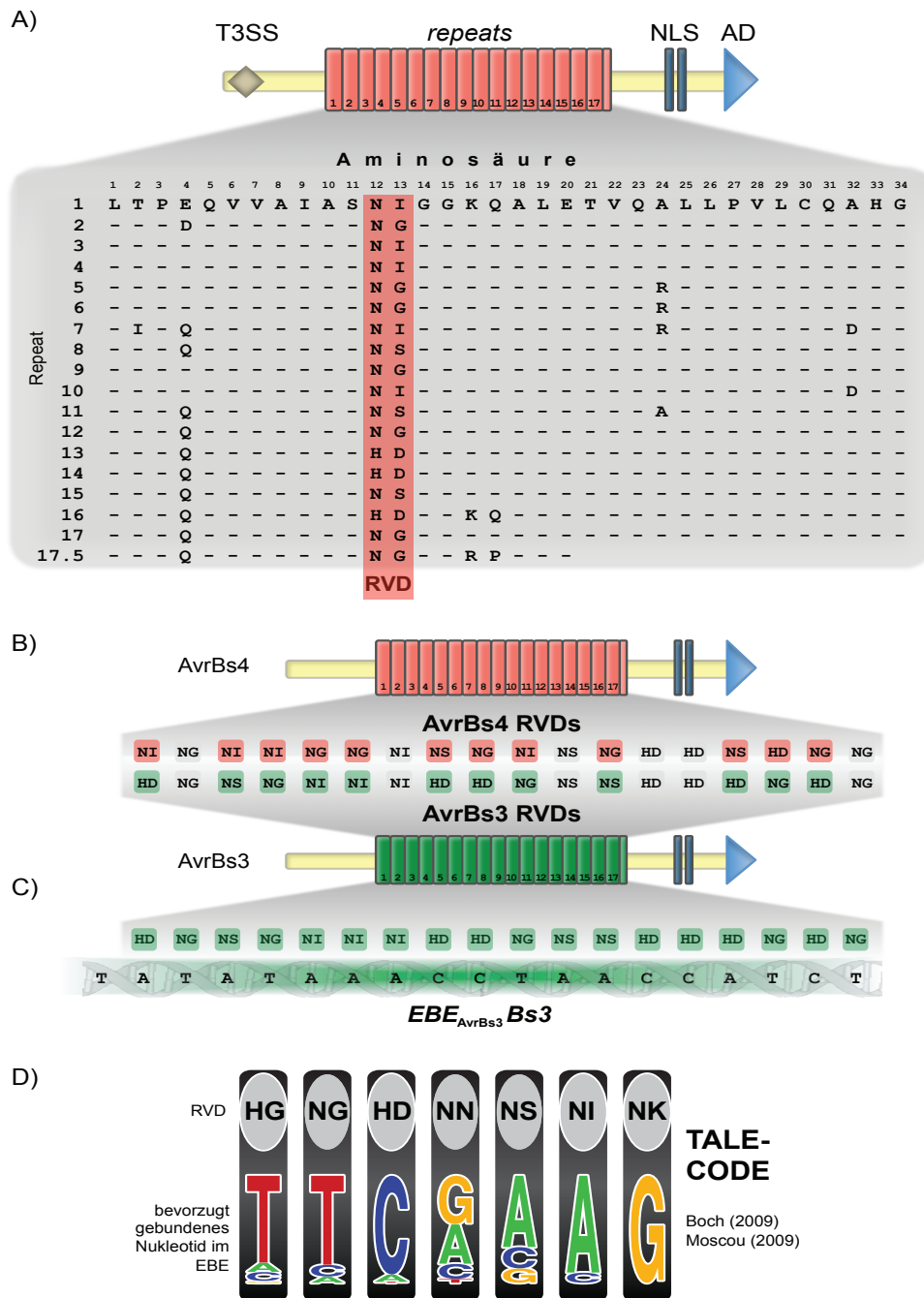
Quelle: <http://www.knowledgebank.irri.org> (<http://creativecommons.org/licenses/by-nc-sa/3.0/>)

Es wurden keine Änderungen vorgenommen.

#### 1.4 *Xanthomonas* TAL-Effektoren – Struktur und TALE-Code

*Xanthomonas*-Bakterien translozieren über ihr Typ-III-Sekretionssystem etwa 30-40 Effektoren in die Pflanzenzelle, die gemeinsam zur Virulenz der Pathogene beitragen (Boch and Bonas, 2010). Eine Gruppe dieser Effektorproteine stellt die Familie der TAL-Effektoren dar, welche nach Translokation in die Wirtszelle in den Zellkern importiert werden und dort Wirtsgene transkriptionell induzieren können (Abb. 3) (Schornack et al., 2013). TAL-Effektoren sind modular aufgebaut und weisen im amino-terminalen (N-Terminus) Bereich ein Typ III-Sekretionssignal auf, das für die Sekretion und Translokation in die Pflanzenzelle notwendig ist (Abb. 2A). Im carboxy-terminalen (C-Terminus) Bereich befinden sich Kernlokalisierungssignale (NLS, *nuclear localization signal*), welche den Transport in den Zellkern vermitteln sowie eine für Transkriptionsfaktoren charakteristische Aktivierungsdomäne (AD, *acidic activation domain*) (Abb. 2A). Im zentralen Bereich befindet sich die sogenannte *repeat*-Region, die aus einer variablen Anzahl von bis zu 33,5 Wiederholungen eines 33 bis 35 Aminosäuremotivs (*repeat*) besteht (Abb. 2A), (Boch and Bonas, 2010). Ein Vergleich der *repeats* eines TAL-Effektors miteinander, ist in Abb. 2A am Beispiel des *Xcv* TAL-Effektors AvrBs4 dargestellt, zeigte, dass sie sich vorwiegend an den Positionen 12 und 13, den sogenannten *repeat variable diresidues* (RVDs) unterscheiden (Boch et al., 2009; Moscou and Bogdanove, 2009). Die RVD-Komposition verschiedener TAL-Effektoren kann sehr unterschiedlich sein wie am Beispiel von AvrBs3 und AvrBs4 in Abb. 2B zu erkennen ist. Die Analyse der RVDs verschiedener *Xanthomonas*-Stämme ergab, dass die RVDs HD, NG, NI, NS und NN am häufigsten vorkommen (Schornack et al., 2013).

Zahlreiche Analysen von TAL-Effektoren und deren induzierten Wirtsgenen haben gezeigt, dass die *repeat*-Region als DNA-Bindedomäne fungiert und die RVDs die Spezifität der DNA-Bindung determinieren (Boch and Bonas, 2010). Es konnte eine Korrelation zwischen den TAL-Effektor RVDs und den Nukleotiden in den Promotoren der Wirtsgene festgestellt werden, die im sogenannten TALE-Code beschrieben ist (Abb. 2C+D) (Boch et al., 2009; Moscou and Bogdanove, 2009). Die am häufigsten vorkommenden RVD-Nukleotid-Kombinationen sind demnach HD:C, NG:T, NI:A und NN:G (RVD:Nukleotid) (Schornack et al., 2013). Die Bindesequenz eines TAL-Effektors im Promotor eines Zielgens wird als Effektorbindeelement (EBE, *effector binding element*) bezeichnet (Abb. 2C) (Boch et al., 2014; Schornack et al., 2013). Ko-Kristallisationsstudien von TAL-Effektoren und ihren korrespondierenden EBE zeigten, dass sie zusammen eine rechtsgängige superhelikale Struktur ausbilden, welche sich um die große Furche der DNA-Doppelhelix windet (Deng et al., 2012; Mak et al., 2012). Dabei besteht jeder *repeat* aus zwei  $\alpha$ -Helices, die über eine kurze Loop-Region, welche die RVDs aufweist, miteinander verbunden sind. Interessanterweise stellte sich auch heraus, dass ausschließlich die Aminosäure an Position 13, der sogenannte *base specifying residue* (BSR), für die Ausbildung der spezifischen Bindung an die DNA verantwortlich ist (de Lange et al., 2014a; Deng et al., 2012; Mak et al., 2012). Analysen zeigten außerdem, dass RVDs, welche an Position 13 identische AS besitzen, gleiche Nukleotide binden. Beispielsweise binden HD und ND beide spezifisch Cytosin- und die RVDs NI und HI beide Adenin-Nukleotide (Cong et al., 2012). Alle anderen Aminosäuren, die sogenannten non-BSR, stabilisieren zum einen die Struktur innerhalb eines *repeats* sowie zwischen verschiedenen *repeats* und gehen unspezifische Bindungen mit dem Zucker-Phosphat-Rückgrad der DNA ein (Mak et al., 2012). Non-BSRs haben zwar keinen Einfluss auf die Spezifität der TAL-DNA-Bindung jedoch konnte ihnen eine Bedeutung auf deren Affinität nachgewiesen werden (Meckler et al., 2013).



**Abb. 2 Struktur von TAL-Effektoren, RVDs und TALE-Code**

A) Struktur eines TAL-Effektors am Beispiel von AvrBs4. Im N-Terminus befindet sich das Typ-III-Sekretionssignal (T3SS) für die Translokation, im C-Terminus Kernlokalisierungssignale (NLS) für den Kernimport und die saure Aktivierungsdomäne (AD) für die transkriptionelle Aktivierung. Die zentrale *repeat*-Region besteht aus 17,5 Wiederholungen eines 34 AS-Motivs und vermittelt die Bindung an die DNA. Die AS-Sequenz (1-34) der *repeats* ist vergleichend zueinander dargestellt. Dabei sind identische AS durch Striche gekennzeichnet und abweichende AS durch ihren Einbuchstaben-Codes. Die *repeats* unterscheiden sich vorwiegend an Position 12 und 13, den sogenannten *repeat variable diresidues* (RVDs). B) Aufgrund der unterschiedlichen RVDs-Komposition von AvrBs4 und AvrBs3, insgesamt 12 voneinander abweichende RVDs, weisen sie unterschiedliche Bindspezifitäten auf. Vergleichende Darstellung der RVDs von den *Xcv* TALEs AvrBs4 und AvrBs3. Identische RVDs sind grau hinterlegt, während abweichende RVDs in der Farbe des jeweiligen TALEs dargestellt sind. C) Darstellung der RVD-Komposition von AvrBs3 mit der Zielsequenz im *Bs3*-Promoter ( $EBE_{AvrBs3}Bs3$ ). D) Der TALE-Code beschreibt die Spezifität der RVDs in Bezug auf die Nukleotide in den EBE von TAL-Zielgenen. Die RVDs HG und NG binden bspw. bevorzugt an ein T-Nukleotid, während HD C-Nukleotide präferiert.



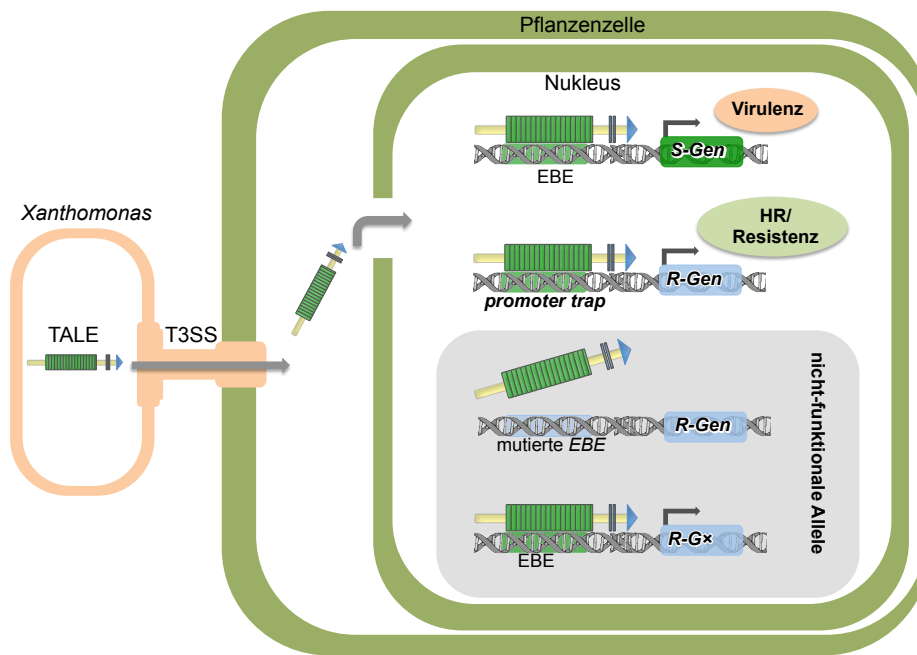
Mit Hilfe des TALE-Codes ist die Vorhersage von EBE (Boch et al., 2014; Römer et al., 2010b; Schornack et al., 2013) und somit die Identifizierung potentieller Zielgene von TAL-Effektoren möglich. Dafür sind bereits verschiedene Programme (Storyteller, Talgetter, Target Finder und Talvez) verfügbar, die eine EBE-Vorhersage *in silico* in kompletten Genom- oder Promoter-Sequenzen zulassen (Doyle et al., 2012; Grau et al., 2013; Perez-Quintero et al., 2013). Des Weiteren können anhand des TALE-Codes auch EBEs generiert werden, welche von korrespondierenden TAL-Effektoren induziert werden (Boch et al., 2009; Hummel et al., 2012; Römer et al., 2009a). Außerdem erlaubt der TALE-Code, TAL-Effektoren mit benutzerdefinierter Erkennungsspezifität, sogenannte designer TALEs (dTALEs) zu erstellen, die korrespondierende Zielgene induzieren können (Boch et al., 2009; Morbitzer et al., 2010). Fusionen von TAL-DNA-Bindeelementen mit Nukleasen, Aktivierungs- oder Repressordomänen wurden bereits erfolgreich zur gezielten Genommodifizierung in verschiedensten Organismen verwendet und stellen ein großes Potential für weitere Applikationen dar (Beispiele finden sich in Bogdanove, 2014; Jankele and Svoboda, 2014; Schornack et al., 2013).

Kürzlich konnten neben RipTALs (*Ralstonia injected protein*) aus *Ralstonia solanacearum* auch Bats (*Burkholderia* TAL-like) aus *Burkholderia rhizoxinica* identifiziert werden, die Ähnlichkeiten zu TAL-Effektoren aus *Xanthomonas* zeigen (de Lange et al., 2013; de Lange et al., 2014b). Während RipTALs ebenso wie TALEs als Transkriptionsaktivatoren agieren, vermittelt die *repeat*-Domäne der Bats ausschließlich Bindspezifität. Obwohl RipTALs und Bats einen größeren Polymorphiegrad verglichen zu TALEs aufweisen, bestimmen weiterhin die BSRs die Bindspezifität.

### 1.5 *Xanthomonas* TAL-Effektoren - Virulenzfunktion

Wie bereits erwähnt induziert *Xanthomonas* Wirtsgene, wobei für einige bereits gezeigt werden konnte, dass ihre TAL-abhängige Induktion das Wachstum der Bakterien begünstigt bzw. für die Ausbildung von Krankheitssymptomen entscheidend sind; weshalb diese Wirtsgene auch als Suszeptibilitätsgene (*S*-Gene) bezeichnet werden (Abb. 3). *OsSWEET11* und *OsSWEET14* aus Reis (*Oryza sativa*), welche für Membranproteine kodieren, sind unter anderem in den Zuckertransport aus dem Phloem in den Apoplasten involviert und begünstigen vermutlich somit die

„Nährstoffversorgung“ des Bakteriums (Chen et al., 2010; Chen, 2014). Erstaunlicherweise kann *OsSWEET14* von verschiedenen *Xanthomonas oryzae* pv. *oryzae* (*Xoo*) TAL-Effektoren (PthXo3, Tal5, TalC, AvrXa7) induziert werden (Antony et al., 2010; Streubel et al., 2013; Yu et al., 2011). Die Induktion des *S*-Gens *CsLOB1* in Zitrus erfolgt ebenfalls durch verschiedene TAL-Effektoren (PthA, PthB, PthC, PthA4 und PthAw) aus *Xanthomonas citri* spp. (Hu et al., 2014; Li et al., 2014). Die Expression von *CsLOB1*, welches für einen Transkriptionsfaktor mit LOB (*lateral organ boundaries*)-Domäne kodiert, aktiviert in der Folge die Expression von Genen, wie bspw.  $\alpha$ -Expansin oder Pectatlyase, die eine Funktion in der Zellwandmorphologie besitzen (Hu et al., 2014). Die TAL-induzierte Expression von *CsLOB1* begünstigt sowohl die Ausbildung von Krankheitssymptomen (Hyperplasie) als auch das Wachstum der Bakterien in der Pflanze (Hu et al., 2014; Li et al., 2014). Ein weiteres potentielles Suszeptibilitätsgen ist *UPA20* (*upregulated by AvrBs3*) in *C. annuum* ECW, das ebenfalls für einen Transkriptionsfaktor kodiert und durch den *Xcv* TAL-Effektor AvrBs3 induziert wird. Die Expression von *UPA20* resultiert in der Aktivierung weiterer Gene, wie *UPA7*, das für ein  $\alpha$ -Expansin kodiert sowie in der Ausbildung einer Hypertrophie von Mesophyllzellen (Kay et al., 2007; Marois et al., 2002). Es wird angenommen, dass die Hyperplasie in Zitrus und die Hypertrophie in Paprika die Freisetzung der Bakterien erleichtert und somit ihre Verbreitung begünstigt (Hu et al., 2014; Kay et al., 2007; Li et al., 2014).



**Abb. 3 Virulenz- und Resistenzinduktion durch TAL-Effektoren in der Pflanze**

TAL-Effektoren werden durch *Xcv* in die Pflanzenzelle transloziert und in den Nukleus importiert. Im Kern induzieren sie Wirtsgene, die zur Virulenz des Pathogens beitragen. Die Induktion erfolgt durch Bindung an EBE im Promoter der Wirtsgene und deren transkriptionelle Aktivierung. In resistenten Pflanzen wird der TALE in eine Falle "gelockt" (*promoter trap*), indem er an ein EBE eines *R*-Genpromoters bindet und diesen aktiviert. Die Expression des *R*-Gens führt zur Ausbildung einer Zelltodreaktion (HR) oder Resistenz, die das Pathogen in seiner Vermehrung und Verbreitung einschränkt bzw. daran hindert. Bei nicht-funktionalen Allelen kann ein mutiertes EBE die Bindung durch den TALE unterbinden und somit die Aktivierung des *R*-Gens. Außerdem können mutierte *R*-Gene (*R-G<sup>x</sup>*) zwar induziert werden, jedoch resultiert ihre Expression aufgrund von Polymorphismen in der kodierenden Sequenz nicht in einer HR/Resistenzreaktion.

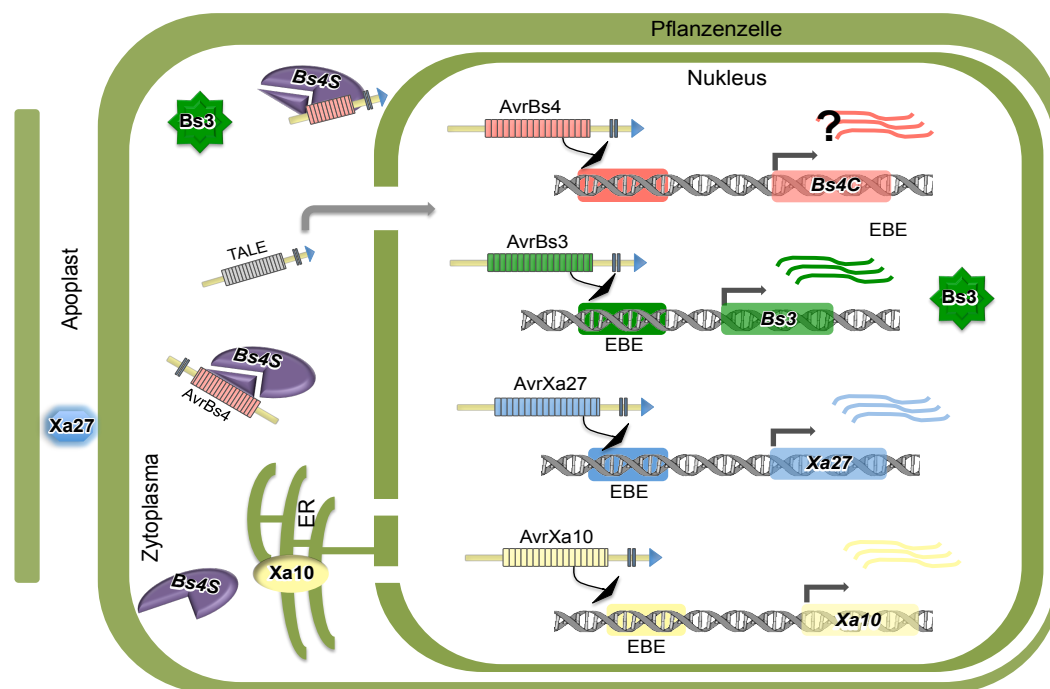
## 1.6 TAL-Effektor R-Proteine - Induktion pflanzlicher Abwehrreaktionen

Die primäre Funktion von TAL-Effektoren ist die Induktion von Wirtsgenen, deren Expression die Pathogenese und Virulenz des Bakteriums begünstigt (siehe 1.5, Abb. 3). Interessanterweise wurden aber viele TAL-Effektoren ursprünglich als Avirulenzdeterminanten identifiziert. So wurde beispielsweise AvrBs3 aus *Xcv* basierend auf seiner Erkennung in resistenten Paprikapflanzen (*C. annuum* ECW-30R), die das *R*-Gen *Bs3* besitzen, isoliert (Bonas et al., 1989; Römer et al., 2007). Es konnte nachgewiesen werden, dass AvrBs3 an ein EBE im *Bs3*-Promoter bindet und diesen transkriptionell aktiviert (Römer et al., 2007). Die susceptible Akzession (*C. annuum* ECW) weist hingegen eine 13-bp Insertion im EBE des Promoters auf und daher wird das *Bs3*-Allel (*Bs3-E*) dieser Linie durch AvrBs3 nicht induziert. Allerdings ist ein AvrBs3-Deletionsderivat (AvrBs3Δrep16), dem die *repeats* 11-14 fehlen, in der Lage, an das EBE von *Bs3-E* zu binden und die Expression von *Bs3-E* zu aktivieren (Römer et al., 2007). Erstaunlicherweise „passen“ die RVD-

Komposition des Deletionsderivat AvrBs3 $\Delta$ rep16 und die korrespondierende Nukleotidsequenz im *Bs3-E*-Promoter (EBE<sub>AvrBs3 $\Delta$ rep16</sub>*Bs3-E*) zusammen, sodass eine Induktion von *Bs3-E* erfolgt. Polymorphismen wurden ebenfalls in den Promotoren von *Xa27*-resistenten und *xa27*-suszeptiblen Reispflanzen nachgewiesen, die bedingen, dass der TAL-Effektor AvrXa27 aus *Xoo* nur den *Xa27*-Promoter, nicht aber den *xa27*-Promoter induzieren kann (Römer et al., 2009a). An den Beispielen *Xa27* und *Bs3* wird deutlich, dass diese Pflanzengenotypen eine sogenannte „promotor-trap“ (Promoterfalle) zur Erkennung von *Xanthomonas* TAL-Effektoren evolviert haben (Abb. 3). Eine Bindung an den Promoter und eine transkriptionelle Induktion der *downstream R*-Gene durch die TALEs resultiert in der Auslösung pflanzlicher Abwehrreaktionen und damit in Resistenz (Abb. 3). Dies stellt einen völlig anderen Erkennungsmechanismus dar, verglichen zu dem der „klassischen“ R-Proteine, bei denen eine direkte oder indirekte Interaktion zur Erkennung des Effektorproteins führt (siehe 1.2.).

TAL-induzierte R-Proteine fungieren als Exekutoren, da sie eine Resistenzreaktion induzieren, die mit der Ausbildung einer HR einhergehen kann und somit das Wachstum und/oder die Verbreitung der Bakterien im Gewebe verhindert wird (Abb. 3). Exekutorgene sollten einer strengen Regulation unterliegen und deshalb war es nicht verwunderlich als nachgewiesen wurde, dass sie nur in Anwesenheit des TALEs transkriptionell induziert werden, in deren Abwesenheit aber transkriptionell „still“ sind (Römer et al., 2007; Schornack et al., 2013). Obwohl *Bs3* und *Xa27*, beide TAL-induziert sind, weisen ihre *R*-Genprodukte keine Homologien zueinander auf. *Bs3* kodiert für eine atypische Flavin-abhängige Monooxygenase (FMO, *flavin monooxygenase*), die im Kern und Zytoplasma lokalisiert ist, während *Xa27* im Apoplasten nachgewiesen wurde und keine Homologie zu Proteinen bekannter Funktion aufweist (Abb. 4)(Gu et al., 2005; Römer et al., 2007; Wu et al., 2008). Kürzlich konnte das TAL-induzierte *R*-Gen *Xa10* aus Reis isoliert werden, welches einen weiteren Exekutor repräsentiert (Tian et al., 2014). Strukturvorhersagen und Lokalisationsstudien zeigten, dass *Xa10* ein Transmembran(TM)-Protein ist, das in der ER-Membran lokalisiert ist (Abb. 4). Eine mögliche Funktion als Calcium-Kanal wird diskutiert (Tian et al., 2014). Diese Arbeiten zeigen, dass TAL-induzierte *R*-Gene für sehr diverse *R*-Genprodukte kodieren und strukturell verschieden von „klassischen“ NBS-LRR-Proteinen sind.

Die einzige bislang bekannte Ausnahme ist das R-Protein Bs4S aus Tomate (*Solanum lycopersicum*), das die Erkennung des TAL-Effektors AvrBs4 aus *Xcv* vermittelt (Ballvora et al., 2001; Bonas et al., 1993). *Bs4S* wird konstitutiv exprimiert und kodiert für ein NBS-LRR-Protein (Schornack et al., 2004). Lokalisationsstudien zeigten, dass Bs4S im Zytoplasma lokalisiert (Abb. 4). Interessanterweise vermittelt Bs4S nicht nur Resistenz gegenüber Stämmen mit AvrBs4-Volllängen-Proteinen sondern auch Derivaten, denen NLS- und AD-Domänen fehlen (Abb. 4) (Schornack et al., 2004). Demnach erfolgt die Bs4-Erkennung durch einen anderen Mechanismus im Vergleich zu allen weiteren TALE-*R*-Genen, für die gezeigt wurde, dass NLS- und AD-Domänen essentiell für die „Erkennung“ sind. Es ist anzunehmen, dass die Perzeption im Zytoplasma durch direkte oder indirekte Interaktion von R-Protein und TAL-Effektor erfolgt (Abb. 4). Damit stellt *Bs4S* ein atypisches TAL-*R*-Gen dar.



**Abb. 4 Übersicht über TAL-*R*-Gen-Interaktionen**

*Xanthomonas* TAL-Effektoren binden an EBE und induzieren die Expression von *R*-Genen, die für sehr diverse R-Proteine kodieren und keine Homologie zueinander aufweisen. AvrBs3 induziert *Bs3*, welches für eine Flavin-abhängige Monooxygenase kodiert und im Zytoplasma sowie im Zellkern lokalisiert, wobei letzteres essentiell für die HR-Induktion ist. *Xa27* wird durch AvrXa27 induziert und sein Genprodukt konnte im Apoplasten nachgewiesen werden. Der TALE AvrXa10 induziert die Expression von *Xa10*, das wahrscheinlich für einen Calcium-Kanal in der Membran des endoplasmatischen Retikulum kodiert. Vermutlich bindet AvrBs4 an das EBE im Promoter des zu isolierenden *R*-Gens *Bs4C* und induziert seine Expression, was zu einer HR führt. AvrBs4 kann außerdem durch das NBS-LRR Bs4S im Zytoplasma erkannt werden. Dies erfolgt durch direkte oder indirekte Interaktion und nicht durch transkriptionelle Aktivierung.

### **1.7 Isolierung von *R*-Genen mittels *map-based cloning approach* (kartengestützten Ansatz)**

Alle bislang bekannten TALE-induzierten *R*-Gene (*Xa10*, *Bs3* und *Xa27*) sowie *Bs4S* wurden in sogenannten kartengestützten Ansätzen (*map-based cloning approaches*) isoliert, die sehr zeitintensiv und technisch aufwendig sind (Peters et al., 2003). So dauerte beispielsweise die Isolierung von *Bs3* von der Initiierung bis zur finalen Publikation mehr als ein Jahrzehnt (Pierre et al., 2000; Römer et al., 2007; Van den Ackerveken et al., 1996). Der kartengestützte Ansatz verwendet neben Methoden der klassischen Genetik auch die der Molekularbiologie. Zuerst erfolgt die genetische Kartierung des Gens, was die Identifizierung gekoppelter molekularer Marker sowie Kopplungsanalysen auf eine Kartierungspopulation erfordert. Anschließend wird eine Markerverdichtung am Ziellocus durchgeführt. Während der nachfolgenden physikalischen Kartierung werden genomische Klone aus Genombibliotheken, wie bspw. YAC (*yeast artificial chromosome*)- oder BAC (*bacterial artificial chromosome*)-Bibliotheken, isoliert. Im letzten Schritt werden Komplementationsanalysen durchgeführt, die Subklonierungen und Infektionsassays *in planta* beinhalten (Burke et al., 1987; Shizuya et al., 1992). Sowohl das Vorhandensein von genomischen Bibliotheken als auch die Verfügbarkeit einer Kartierungspopulation sind Voraussetzung für den Ansatz.

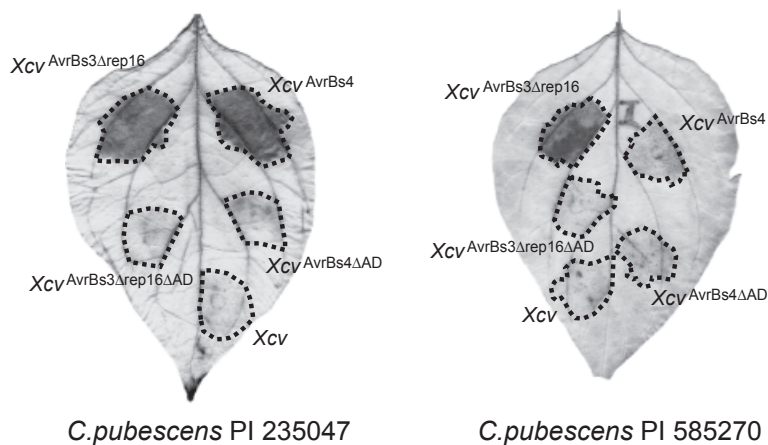
Die Isolierung der TAL-induzierten *R*-Gene erfolgte mittels *map-based cloning approach*, bevor bekannt war, dass diese in Abwesenheit des TALEs nicht exprimiert sind (Römer et al., 2007, Schornack et al., 2013) und erst durch diese transkriptionell induziert werden. Für die Isolierung weiterer TAL-*R*-Gene würde sich aber nun ein alternativer Ansatz, nämlich eine Identifizierung über Transkriptomanalysen anbieten. Der Vergleich der Transkriptomdaten einer resistenten und einer suszeptiblen Akzession nach Infektion mit *Xanthomonas*-Stämmen, welche die entsprechenden TALEs exprimieren, sollte das *R*-Gen identifizieren.

### **1.8 Arbeitshypothese: Isolierung eines TAL-induzierten *R*-Gens mittels RNA-Seq-basierter Transkriptomanalysen**

Wie bereits angeführt, gehört AvrBs4 ebenfalls zu den TAL-Effektoren aus *Xcv* (siehe 1.4). Es war bekannt, dass AvrBs4 nicht nur in Tomate (siehe 1.6) sondern auch in Paprika *Capsicum pubescens* (*C. pubescens*, PI 235047) erkannt wird und eine Resistenzreaktionen auslöst (Ballvora et al., 2001; Minsavage et al., 1999; Stall

et al., 2009). Ziel dieser Arbeit war die Isolierung des *R*-Gens *Bs4C* aus *C. pubescens* mittels RNA-Seq-basierter Transkriptomanalyse. Um zu prüfen, ob der Ansatz für *Bs4C* geeignet ist, wurden AvrBs4-Deletionsderivate in *C. pubescens* getestet. Ein AvrBs4-Derivat ohne AD-Domäne induzierte keine HR mehr in der Akzession PI 235047 (Abb. 5) (Gürlebeck, 2001; Strauß et al., 2012), was vermuten ließ, dass AvrBs4 als Transkriptionsaktivator in *C. pubescens* fungiert und die Expression seines korrespondierenden Resistenzgens *Bs4C* induziert. In der *C. pubescens*-Akzession PI 585270 induzierten weder AvrBs4-Volllängen-Proteine noch AvrBs4-Derivate eine HR (Abb. 5) (Gürlebeck, 2001; Minsavage et al., 1999; Strauß et al., 2012). Vermutlich ist aufgrund von Promoterpolymorphismen diese Linie nicht mehr durch AvrBs4 induzierbar und somit AvrBs4-suszeptibel. Beide Akzessionen sind somit, angesichts ihrer unterschiedlichen AvrBs4-Reaktion, für einen „proof-of-principle“ Ansatz zur Isolierung der *R*-Gens *Bs4C* mittels RNA-Seq-Analysen geeignet. Durch Vergleiche der Transkripte der suszeptiblen und resistenten Akzession, welche mittels *next generation sequencing* (NGS)-Technologien sequenziert werden, sollte es möglich sein, *Bs4C* zu isolieren.

Für den unwahrscheinlichen Fall, dass das *R*-Gen in beiden Akzessionen induziert wird aber aufgrund eines Polymorphismus in der suszeptiblen Akzession nicht funktional ist, könnte dies auch identifiziert werden (Abb. 3). Die Transkriptzahlen für das Kandidatengen sollten dann in beiden Akzessionen entsprechend ähnlich sein und in beiden Akzessionen ein funktionales EBE für den TAL-Effektor identifiziert werden können. Die funktionale Analyse der Exekutor-Allele würde mittels Überexpressionsanalyse (in *N. benthamiana*) erfolgen und nur für das Allel aus der resistenten Akzession zur HR-Induktion führen. Jedoch ist solch ein Fall nicht bekannt und alle bislang isolierten TAL-*R*-Gen-Allele aus suszeptiblen und resistenten Akzessionen zeigten funktional-relevante Polymorphismen im EBE (Abb. 3, Schornack et al., 2013).



**Abb. 5 Die Erkennung von AvrBs4 in *C. pubescens* PI 235047 durch *Bs4C* ist abhängig von der Aktivierungsdomäne.**

Der WT *Xanthomonas*-Stamm 85-10 (*Xcv*) und *Xanthomonaden*, die AvrBs4 (*Xcv*<sup>AvrBs4</sup>) und AvrBs3rep16 (*Xcv*<sup>AvrBs3<sub>Δ</sub>rep16</sup>) exprimieren sowie Derivate, bei denen die Aktivierungsdomäne (AD) (*Xcv*<sup>AvrBs4<sub>Δ</sub>AD</sup> und *Xcv*<sup>AvrBs3<sub>Δ</sub>rep16<sub>Δ</sub>AD</sup>) deletiert ist, wurden mittels kanülenloser Spritze in die Blätter der *C. pubescens* Akzessionen PI 235047 und PI 585270 inokuliert. Alle *Xanthomonas*-Stämme wurden auf eine Bakteriendichte von  $5 \times 10^8$  cfu/ml eingestellt. Die infiltrierten Blattbereiche wurden durch gestrichelte Linien markiert. Die Blätter wurden nach drei Tagen geerntet und zum Hervorheben der HR in Ethanol entfärbt. AvrBs3rep16 (*Xcv*<sup>AvrBs3<sub>Δ</sub>rep16</sup>)-exprimierende *Xanthomonaden* dienten als Kontrolle da gezeigt werden konnte, dass er das *C. pubescens* *Bs3*-Allel in AD-abhängiger Weise aktiviert (Strauß 2012, Promotion Patrick Römer).

Abbildung modifiziert nach Strauß et al., 2012 Fig.S4.

## 1.9 Zielstellung der Arbeit

Ziel dieser Arbeit war es, dass *R*-Gen *Bs4C* mittels RNA-Seq-Analysen zu isolieren. Transkriptanalysen zwischen der AvrBs4-resistenten (PI 235047) und der AvrBs4-suszeptiblen (PI 585270) *C. pubescens*-Akzession sollten Kandidatengene ermittelt, die nachfolgend analysiert werden.



## 2 Material und Methoden

### 2.1 Chemikalien, Geräte und Verbrauchsmaterialien

Wenn nicht anders beschrieben, wurden Chemikalien, Geräte und Verbrauchsmaterialien von folgenden Firmen bezogen: Agilent, Analytik Jena (Jena); Applichem (Darmstadt); Bio-Rad Laboratories GmbH (München); Biozym (Hessisch Oldendorf); Carl Roth GmbH & Co. (Karlsruhe); Clontech (Takara Bio Company) (Saint-Germain-en-Laye, Frankreich); Eppendorf (Hamburg); Eurofins (Ebersberg); GE Healthcare (München); Greiner Labortechnik GmbH (Solingen); Life Technologies (Groningen, Niederlande); Merck (Darmstadt); Metabion GmbH (Martinsried); Nanotemper (München); New England Biolabs GmbH (Frankfurt/Main); peqlab Biotechnologie GmbH (Erlangen); Qiagen GmbH (Hilden); Roche Diagnostics GmbH (Mannheim); Sigma-Aldrich Chemie GmbH (Taufkirchen); Thermo Fisher Scientific (Waltham, USA); VWR (Darmstadt).

### 2.2 Verwendete Kits

SMARTer RACE cDNA kit + GenomeWalker – Clontech

GeneJET Plasmid Miniprep Kit (Plasmidisolierung) – Thermoscientific

RNeasy Plant Miniprep/Midiprep kit (RNA-Isolierung) (Qiagen)

RevertAid First Strand cDNA Synthesis Kit (cDNA Synthese Kit) – Thermoscientific

### 2.3 Verwendete Nährmedien und deren Zusammensetzung

Bezeichnung	Zusammensetzung	Verwendung
LB Medium	1 % (w/v) Bakto-Trypton; 1 % (w/v) NaCl; 0,5 % (w/v) Hefe-Extrakt pH 7,5	Anzucht von <i>Escherichia coli</i> ( <i>E.</i>
YEB Medium	0,5 % (w/v) Bakto- <i>beef</i> -Extrakt; 0,1 % (w/v) Bakto-Hefe-Extrakt; 0,5 % (w/v) Sucrose; 0,5 % (w/v) Bakto-Pepton; 0,2 % 1 M MgSO <sub>4</sub> pH 7,2	Anzucht von <i>Agrobacterium tumefaciens</i> ( <i>A. tumefaciens</i> )
NYG Medium	0,5 % (w/v) Bakto-Pepton; 0,3 % (w/v) Hefe-Extrakt; 0,2 % Glycerin, pH 7,2	Anzucht von <i>Xanthomonas campestris</i> pv.
SOC-Medium	2 % (w/v) Bakto-Trypton; 0,5 % (w/v) Bakto-Hefe-Extrakt; 0,05 % (w/v) NaCl; 20 mM Glukose	Anzucht von <i>E. coli</i>

MgSO<sub>4</sub>, Magnesiumsulfat; NaCl, Natriumchlorid

Für die Herstellung von LB-, NYG- und YEB-Agarplatten wurde dem Flüssigmedium je 1,5 % (w/v) Bakto-Agar vor dem Autoklavieren zugegeben. Die Sterilisation der Nährmedien erfolgte durch Autoklavieren für 20 min bei 121°C und  $1,1 \times 10^5$  Pa. Die Lagerung der Agarplatten erfolgte bei 4°C. Nach dem Autoklavieren erfolgte die Zugabe von Antibiotika gemäß der Verwendung für die selektive Kultivierung.

## 2.4 Verwendete Lösungen und Puffer

Bezeichnung	Zusammensetzung
10 x PCR-Puffer	0,1 M Tris-HCl pH 8,5; 0,5 M KCl; 20 mM MgCl <sub>2</sub> ; 1 % (v/v) TritonX 100; 0,1 % (w/v) Gelatine
Ladepuffer	15 % Ficoll 400, Farbstoff: Orange G in TE Puffer pH 7,5
dNTPs	2 mM
4 x Laemmli-Puffer	250 mM Tris-HCl; 40 % Glycerol; 8 % SDS; 20 % $\beta$ -Mercaptoethanol; Bromphenolblau
TAE	40 mM Tris-Acetat pH 8,0; 1 mM EDTA
10 x TBE	90 mM Borsäure; 2 mM EDTA; 90 mM Tris-HCl
DEPC-Wasser	2 % Diethylpyrocarbonat (DEPC) in H <sub>2</sub> O
1 x TE	10 mM Tris-HCl pH 8,0; 0,1 M EDTA
Agrobakterium-Infiltrationsmedium (AIM)	10 mM MgCl <sub>2</sub> ; 5 mM MES pH 5,3; 150 $\mu$ M Acetosyringon
Xcv-Infiltrationsmedium	10 mM MgCl <sub>2</sub>
X-Gal-Lösung	20 mg X-Gal (5-Brom-4-chlor-3-indoxyl- $\beta$ -D-galactopyranosid) in 1 ml Dimethylformamid, Lagerung lichtgeschützt; Arbeitsverdünnung 1:400
GUS-Färbelösung	10 mM Na-Phosphat pH7, 10 mM EDTA pH8, 1 mM Kalium-Ferricyanat, 1 mM Kalium-Ferrocyanat 0,1% Triton-X 100 dH <sub>2</sub> O frisch dazu 0,1% X-Gluc (in DMSO)

## 2.5 Verwendete Größenstandards

Größenstandards	Größen	Hersteller
GeneRuler 1kb DNA Ladder (in kb)	0,25; 0,5; 0,75; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0; 5,0; 6,0; 8,0; 10,0	Thermoscientific
GeneRuler 50bp DNA Ladder plus (in bp)	50, 100, 150, 200, 250, 300, 400, 500, 600, 700, 800, 900, 1000	Thermoscientific

## 2.6 Verwendete Antibiotika (Stammkonzentration und Verdünnungen)

Antibiotika	Abkürzung	Stammkonzentration	gelöst in	Verdünnung	
				Platten	flüssig
Ampicillin	Amp	100 mg/ml	H <sub>2</sub> O	1:1000	1:2000
Chloramphenicol	Cm	15 mg/ml	Ethanol	1:1000	1:1000
Gentamycin	Gent	15 mg/ml	H <sub>2</sub> O	1:1000 1:300*	1:1000
Hygromycin	Hyg	150g/l	DMF	1:500	1:500
Kanamycin	Kan	25 mg/ml	H <sub>2</sub> O	1:1000 1:250*	1:1000
Rifampicin	Rif	100 mg/ml	DMF	1:1000 1:2000*	1:1000
Spectinomycin	Spec	100 mg/ml	H <sub>2</sub> O	1:1000	1:1000
Tetracycline	Tet	10 mg/ml	Ethanol 70%	1:1000	1:1000
* kennzeichnet Antibiotikaverdünnung für <i>Xanthomonas</i> und <i>Agrobacterium</i>					

## 2.7 Verwendete Vektoren

Tab. Verwendete Vektoren			
Vektoren	Resistenz	Beschreibung	Referenz
pDONR222 <sup>TM</sup>	Kan	Bibliotheksausgangsvektor; mit <i>attP</i> -Kassette ( <i>attP1-ccdB-Cm-attP2</i> ) für Erstellung der <i>C. pubescens</i> cDNA-Bibliothek	Invitrogen
pJET1.2	Amp	Klonierung von PCR-Produkten der RACE und GenomeWalker-Analysen	Thermoscientific
pUC57	Amp	Klonierung	Genscript
pUC57ΔBsaI	Amp	Klonierung	(Morbitzer et al., 2011)
pUC57 Gent	Gent	Klonierung	Robert Morbitzer
pUC57 Spec	Spec	Klonierung	Robert Morbitzer
pENTR-BsaI	Kan	Klonierung	Tom Schreiber

Der Vektor pENTR/D Topo wurde so modifiziert, dass die GW-Kassette von BsaI-sites flankiert ist und für cut-ligations verwendet werden kann. Dabei produziert BsaI einen CACC-Überhang und kann für gerichtete Klonierung verwendet und effizient selektiert werden, da nach Transformation nur Klone angezogen werden können, die das Insert anstatt des *ccdB* Gen aufweisen.

## 2.8 Zusammenstellung verwendeter Bakterienstämme

Stamm	Resistenz	Beschreibung (Referenz)
<i>Escherichia coli</i> ( <i>E. coli</i> )		
<i>ccdB</i> survival (One shot <i>ccdB</i> )		F- <i>mcrA</i> Δ( <i>mrr-hsdRMS-mcrBC</i> )Φ80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 ara</i> Δ139 Δ( <i>ara-leu</i> )7697 <i>galU</i>

Survival T1 Phage-resistant cells)		<i>galK rpsL</i> (StrR) <i>endA1 nupG fhuA::IS2</i> (Invitrogen)
DB3.1		F- <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 recA1 ara</i> $\Delta$ 139 $\Delta$ ( <i>ara-leu</i> )7697 <i>galU galK rpsL</i> (StrR) <i>endA1 nupG fhuA::IS2</i> (Invitrogen)
<i>OneShot</i> (TOP10)		F- <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\Phi$ 80 <i>lacZ</i> $\Delta$ M15 $\Delta$ <i>lacX74 recA1 ara</i> $\Delta$ 139 $\Delta$ ( <i>ara-leu</i> )7697 <i>galU galK rpsL endA1 nupG</i> (Invitrogen)
<i>Agrobacterium tumefaciens</i> ( <i>A. tumefaciens</i> )		
GV3101	Rif	Ti-Plasmid: pMP90 (pTiC58 $\Delta$ T-DNA), Gene für Nopalinsynthese (Van Larebeke et al., 1974)
<i>Xanthomonas campestris</i> pv. <i>vesicatoria</i> ( <i>Xcv</i> )		
<i>Xcv</i> 85-10	Rif	( <i>avrBs1</i> <sup>+</sup> , <i>avrBs2</i> <sup>+</sup> , <i>avrBs3</i> <sup>+</sup> )
<i>Xcv</i> 85-10 ( <i>avrBs3</i> )	Rif Spec	( <i>avrBs1</i> <sup>+</sup> , <i>avrBs2</i> <sup>+</sup> , <i>avrBs3</i> <sup>+</sup> )
<i>Xcv</i> 85-10 ( <i>avrBs3</i> $\Delta$ <i>rep16</i> )	Rif Spec	( <i>avrBs1</i> <sup>+</sup> , <i>avrBs2</i> <sup>+</sup> , <i>avrBs3</i> $\Delta$ <i>rep16</i> <sup>+</sup> )
<i>Xcv</i> 85-10 ( <i>avrBs4</i> )	Rif Spec	( <i>avrBs1</i> <sup>+</sup> , <i>avrBs2</i> <sup>+</sup> , <i>avrBs3</i> <sup>+</sup> , <i>avrBs4</i> <sup>+</sup> )

## 2.9 Zusammenstellung verwendeter Oligonukleotide

Eine Zusammenstellung der verwendeten Oligonukleotide und ihre Sequenz befindet sich im Anhang unter A34.

## 2.10 Pflanzenmaterial

Transiente Expression von *A. tumefaciens* in *Nicotiana benthamiana* (*N. benthamiana*) Pflanzen für phänotypische und GUS-Analysen. Alle Pflanzen wurden unter Gewächshausbedingungen angezogen (22-18 °C, 60-70% relative Luftfeuchte, Lichtperiode 16 Stunden).

Phänotypische Analysen der Resistenzreaktion von *Bs4C-R* (oder *Bs4C-R*-Allelen) mit *AvrBs4* wurden in *C. pubescens*-Akzessionen PI 235047, PI 585270, sowie F1- und F2-Nachkommen (PI 235047xPI 585270), weitere *Capsicum*-Akzessionen sowie verschiedenen Nachtschattengewächsen (*Solanaceae*-Spezies) und *Solanum lycopersicum* (*S. lycopersicum*) durchgeführt. Alle Pflanzen wurden unter Gewächshausbedingungen angezogen (22-18 °C, 60-70% relative Luftfeuchte, Lichtperiode 16 Stunden).

## 2.11 Methoden

### 2.11.1 Bakterien

#### 2.11.1.1 Kultivierung von Bakterien

##### *Escherichia coli*

Die Anzucht von *E. coli* Zellen erfolgte entweder in flüssigem LB-Medium mit den entsprechenden Antibiotikazusätzen unter Schütteln in einem Kulturröhrchen oder auf antibiotikahaltigen LB-Platten über Nacht bei 37 °C.

##### *Xanthomonas campestris* pv. *vesicatoria*

Die Anzucht von *Xcv* erfolgte auf NYG-Platten mit den entsprechenden Antibiotikazusätzen über zwei Nächte bei 30 °C. Dann würde etwas Bakterienkultur abgenommen und in flüssigem NYG-Medium gegeben und mit den entsprechenden Antibiotikazusätzen unter Schütteln in einem Erlenmeyerkolben über Nacht bei 30 °C angezogen. Alternativ wurden Bakterien NYG-Platten mit den entsprechenden Antibiotikazusätzen über zwei Nächte bei 30 °C angezogen und in 1 oder 10 mM MgCl<sub>2</sub> resuspendiert.

##### *Agrobacterium tumefaciens*

Die Anzucht von *A. tumefaciens* erfolgte entweder in flüssigem YEB-Medium mit den entsprechenden Antibiotikazusätzen unter Schütteln in einem Kulturröhrchen oder auf antibiotikahaltigen YEB-Platten über zwei Nächte bei 30 °C.

#### 2.11.1.2 Transformation chemisch-kompetenter Bakterien

Von der Plasmidlösung oder dem Ligationsansatz wurden maximal 5 µl zu 50 µl chemisch kompetenten Zellen, die vorher langsam auf Eis aufgetaut worden, gegeben. Der Ansatz wurde vorsichtig durch Rühren gemischt und für 20 Min auf Eis inkubiert. Danach wurde der Ansatz für 45 sec einer Temperatur von 42 °C (Wasserbad) ausgesetzt. Durch den erfolgten Hitzeschock öffnen sich die Zellporen und die Zellen können DNA aufnehmen. Anschließend wurde der Ansatz 2 Minuten auf Eis inkubiert und mit 250 µl SOC-Medium (Invitrogen) oder LB-Medium versetzt. Nach einer Inkubation von 60 min bei 37 °C (optional unter Schütteln) wurde der Ansatz auf LB-Platten mit entsprechenden Antibiotika ausplattiert und über Nacht bei 37 °C angezogen.

#### 2.11.1.3 Transfer von Plasmid-DNA in elektrokompetente Zellen

Es wurden elektrokompetente Bakterien (vorher auf Eis aufgetaut) und *A. tumefaciens*-Stamm GV3101 (vorher auf Eis aufgetaut) für die Transformation verwendet. Nach Zugabe von 50-300 ng DNA wurde der Ansatz gemischt und in eine 2 mm Elektroporationsküvette gegeben. Die Transformation erfolgte bei  $U=2500$  V und  $R=200$   $\Omega$ . Nach Zugabe von 250  $\mu$ l flüssigem LB-Medium (YEB-Medium bei *A. tumefaciens*) und 60 min Inkubation bei 37 °C (30 °C bei *A. tumefaciens*) wurde der Ansatz auf antibiotikahaltigen selektiven Agarplatten ausplattiert und bei 37 °C über Nacht (30 °C für 48 Stunden bei *A. tumefaciens*) angezogen.

#### 2.11.1.4 Dauerhafte Lagerung von Bakterien

Für die dauerhafte Aufbewahrung der Bakterien wurden DMSO-Stammproben erstellt. Dazu erfolgte die selektive Anzucht der Bakterien entweder in einer Flüssigkultur bzw. auf einer Agarplatte. Von der Übernachtskultur wurden 930  $\mu$ l mit 70  $\mu$ l DMSO gemischt. Erfolgte die Anzucht auf Agarplatten, wurden Bakterien mit einem Holzstäbchen abgenommen und in Flüssigmedium mit 7 % (v/v) DMSO resuspendiert. Nach dem Schockgefrieren der Bakterien in flüssigem Stickstoff wurden sie bei -80 °C gelagert.

#### 2.11.1.5 Isolation von Plasmid-DNA aus Bakterien

Die Isolierung von Plasmid-DNA aus *E. coli* erfolgte nach dem Prinzip der alkalischen Lyse (Sambrook et al., 1989). Die Plasmid-DNA wurde unter Verwendung des GeneJet Plasmid MiniPrep Kit (Fermentas) nach Herstellerprotokoll aus 4 ml einer Übernachtskultur isoliert und mit 50-70  $\mu$ l aqua dest. eluiert.

#### 2.11.1.6 Transiente Expression in *N. benthamiana* durch *A. tumefaciens*

Um die Funktionalität der erstellten Vektorkonstrukten zu testen, kann man sie transient im Blattgewebe exprimieren. Dazu werden *A. tumefaciens*-Stämme, die das gewünschte Plasmid enthalten in *N. benthamiana* infiltriert. Die Agrobakterienstämme werden dazu in Flüssigkulturen mit entsprechenden Antibiotika (4 ml) angezogen, abzentrifugiert und das Pellet in Agroinfiltrationsmedium (AIM) resuspendiert. Die Bakterienlösung wurde auf eine optische Dichte ( $OD_{600}$ ) von 0,8 (etwa  $8 \times 10^8$  Zellen/ml) eingestellt und entsprechend gemischt. Die Infiltration der Bakterienlösung erfolgte 1-2 Stunden nach Resuspension in AIM mit einer kanülenlosen Spritze in die Interzellularräume des Blattgewebes der Blattunterseite.

Die phänotypische Auswertung erfolgte 3 – 5 Tage nach Infiltration. Bei HR-Ausbildungen wurden die Blätter geerntet und in Ethanol entfärbt um die HR besser zu visualisieren.

#### 2.11.1.7 GUS-Analysen nach transienter Expression in *N. benthamiana*

Nach Infiltration der Promoter-*uidA*-GUS-Konstrukte mit den jeweiligen TAL-Effektor-Konstrukten wurde zwischen 26-48 Stunden nach Infiltration Blattscheiben ausgestanzt und in einem 2 ml Reaktionsgefäß (Eppendorf Safelock) gesammelt. Nach Zugabe der GUS-Färbelösung wurden mit Hilfe einer Spritze die Luft aus den Blattscheiben durch Anlegen von Vakuum entfernt. Damit waren die Blattscheiben komplett mit der GUS-Färbelösung infiltriert und verblieben dort für 1-2 Tage (optional bei 37°C). Anschließend wurde die GUS-Färbelösung gegen Ethanol getauscht um das Chlorophyll aus den Blattscheiben herauszulösen und somit die Färbung besser sichtbar zu machen. Die Blattscheiben wurden nach dem Entfärben dokumentiert.

#### 2.11.1.8 Inokulationsexperimente mit *Xcv*

Die Analyse der *C. pubescens* Pflanzen der Akzessionen PI 235047 und PI 585270 sowie der F<sub>2</sub>-Generation, die aus der Kreuzung beider Akzessionen hervorging, und verschiedener *Capsicum* und *Solanaceae*-Spezies erfolgte mit verschiedenen *Xcv*-Stämmen. Die Bakterien wurden dafür auf selektiven NYG-Platten für zwei Tage bei 30 °C angezogen. Für die Infiltration wurden Bakterien von der Platte in einer 10 mM MgCl<sub>2</sub>-Lösung suspendiert. Anschließend erfolgte die Einstellung der optischen Dichte der Bakteriensuspension auf 0,4 bei einer Wellenlänge von 600 nm. Das entspricht einer ungefähren Bakterienzahl von  $4 \times 10^8$  Zellen pro ml. Mit einer kanülenlosen Spritze wurde die Bakteriensuspension vorsichtig in die Interzellularräume der Blattunterseite junger Blätter injiziert. Die Auswertung der Phänotypen erfolgte 3-5 Tage nach der Infiltration.

Für die RNA-Isolierung (Transkriptanalysen) wurde das Blattmaterial nach 6, 12, 18 und 24 Stunden nach Infiltration geerntet und in flüssigem Stickstoff schockgefrostet. Desweiteren wurde nicht-inokuliertes Material geerntet, welches als Kontrolle diente. Für die RNA-Isolierung (sq RT-PCR, Zeitkurve) wurde das Blattmaterial 3, 6, 9, 12, 15, 18, 21 und 24 Stunden nach Infiltration geerntet und in flüssigem Stickstoff schockgefrostet. Auch hier diente nicht-inokuliertes Material als Kontrolle

## 2.12 DNA-Methoden

### 2.12.1 Agarose-Gelelektrophorese

Um PCR-amplifizierte Fragmente oder durch Restriktionsendonukleasen gespaltene DNA-Fragmente ihrer Größe nach aufzutrennen oder DNA und RNA auf ihre Qualität zu untersuchen oder ihre Menge abzuschätzen, verwendet man die Agarose-Gelelektrophorese. Nach dem Prinzip der Elektrophorese „wandern“ die negativ geladenen DNA-Fragmente entsprechend ihrer Größe unterschiedlich schnell im Agarosegel in Richtung Pluspol. Die Agarose kann unterschiedlich stark konzentriert sein. Niederprozentige Gele (1 %ig) eignen sich gut zur Auftrennung von großen Fragmenten, höherprozentige Gele (3 %ig) werden eher für kurze Fragmente bevorzugt. Für die Herstellung der Agarose wird die Menge eingewogen und in dem entsprechenden Volumen 1 x TAE / 1 x TBE (bei RNA) gelöst und in der Mikrowelle kurz aufgeköcht. Nach dem Abkühlen der Lösung wird diese mit Ethidiumbromid versetzt (3 µl pro 100 ml Agaroselösung) und kann in Horizontalgelapparaturen gegossen. Die aufzutragenden Proben werden mit DNA-Ladepuffer versetzt und können nach Aushärten des Gels aufgetragen werden. Die Auftrennung der Fragmente erfolgt bei 90 – 150 Volt in 1 x TAE als Laufpuffer. Das zugegebene Ethidiumbromid dient zur Visualisierung der DNA, da es in sie interkaliert und durch UV-Licht sichtbar wird. Zur Abschätzung der Größen der DNA-Fragmente wurde ein Größenstandard (siehe 2.5) mit aufgetragen werden.

### 2.12.2 Polymerasekettenreaktion (PCR)

Um DNA-Sequenzen selektiv zu vervielfältigen bedient man sich dem Verfahren der PCR (Saiki et al., 1988). Dabei werden mehrere Zyklen eines bestimmten Programms durchlaufen, die es der DNA-Polymerase (0,2 µl) im entsprechenden Polymerase-Puffer ermöglichen, mehrere neue DNA-Doppelstränge ausgehend von einer Matrize zu synthetisieren. Die DNA-Polymerase benötigt neben Desoxynukleosidtriphosphaten (200 µM dNTP), kurze Oligonukleotide (Primer, 300 mM), die als Startmolekül dienen und sich komplementär an den zu amplifizierenden Nukleinsäurebereich anlagern. Ein PCR-Zyklus besteht aus 3 Phasen. Die Denaturierung der doppelsträngigen DNA wird gefolgt von der Anlagerung der Oligonukleotide und der Synthese des komplementären Stranges durch Anknüpfen von dNTPs an die vorher gebundenen Oligonukleotide durch die DNA-Polymerase. Die mehrfache Wiederholung der drei Phasen (Denaturierung, Primerbindung,



Elongation) führt zur exponentiellen Vervielfältigung des Fragments. Die verwendete DNA-Polymerase ist in der Regel, die hitzestabile *Taq*-DNA-Polymerase, die aus dem thermophilen Bakterium *Thermus aquaticus* isoliert wurde, und welche eine hohe Optimierungstemperatur ( $>70^{\circ}\text{C}$ ) aufweist und auch bei sehr hohen Temperaturen ( $95^{\circ}\text{C}$ ) nicht denaturiert wird. Die *Taq*-Polymerase besitzt eine sehr hohe Prozessivität aber keine Korrekturlesefunktion (*proof reading*). Dadurch kann es sein, dass auch nicht zum Matrizenstrang komplementäre Nukleotide eingebaut werden können. Es gibt aber auch DNA-Polymerasen mit Korrekturlesefunktion. Ein Beispiel hierfür ist die *Phusion*-Polymerase, meist dann verwendet, wenn fehlerfreie PCR-Produkte benötigt werden. Die PCR-Reaktionen wurde in Temperaturzyklus-Steuergeräten (Thermocycler) mit beheizbarem Deckel verschiedener Hersteller durchgeführt.

PCR-Programm:  $96^{\circ}\text{C}/2\text{min} - [ 96^{\circ}\text{C}/10\text{sec} - 50\text{-}70^{\circ}\text{C}/20\text{sec} - 72^{\circ}\text{C}/30\text{sec-}4\text{min} ] \times 30\text{-}35 - 72^{\circ}\text{C}/2\text{min} - 4^{\circ}\text{C}$

#### 2.12.2.1 Kolonie-PCR

Die Kolonie-PCR ermöglicht die Untersuchung einer großen Anzahl von Bakterienklonen auf die An-/Abwesenheit eines DNA-Fragmentes. Dabei kann auf eine vorherige DNA-Extraktion der Klone verzichtet werden. Nachdem die Reaktionsansätze pipettiert worden sind, nimmt man mit einer Spitze etwas von der Bakterienkolonie auf, streicht es auf einen neuen Platte über und suspendiert den Rest im Reaktionsansatz. Der Aufschluss des Zellmaterials erfolgt durch einen verlängerten, ersten Denaturierungsschritt (5 min bei  $95^{\circ}\text{C}$ ). Daran schließt sich ein Standard-PCR-Programm an.

#### 2.12.3 Restriktionsanalysen

Für die sequenzspezifische Spaltung von PCR-Produkten bzw. Plasmid- und genomischer DNA wurden Restriktionsendonukleasen der Firmen Thermoscientific (Fermentas) oder New England Biolabs (NEB) verwendet. Es wurden die für die einzelnen Restriktionsendonukleasen, laut Firma, angegebenen Puffer und Inkubationstemperaturen verwendet. Die Inaktivierung der Enzyme erfolgte mittels Hitzeinaktivierung nach Angaben der Hersteller.

#### 2.12.4 GATEWAY-basierte Klonierung von DNA-Fragmenten

Das GATEWAY-System basiert auf dem Rekombinationssystem des Bakteriophagen  $\lambda$ . Dem Phagen ist es möglich seine DNA in die Bakterienzelle zu injizieren und in das Genom des Wirtes zu integrieren. Dies erfolgt durch die sequenzspezifische (*site-specific*) Rekombination homologer DNA-Sequenzen (*attachment sites*, *att*-Stellen) zwischen dem Phagen (*attP*-Stellen) und dem Bakterium (*attB*-Stellen). Dabei entstehen *attL*- und *attR*-Stellen, die sogenannten Hybridsequenzen von *attB* und *attP* darstellen. Das Ausschneiden des Phagengenoms erfolgt über die Rekombination zwischen *attL*- und *attR*-Stellen. Die GATEWAY-Technologie basiert auf den LR- und den BP-Rekombinationsreaktionen. Bei der LR-Rekombination erfolgt die Rekombination zwischen dem Entry-Klon, dessen DNA-Fragment von *attL*-Stellen flankiert ist und einem Destinationsvektor, welcher einen negativen Selektionsmarker enthält, der von *attR*-Stellen flankiert ist. Der entstandene Expressionsvektor trägt das von *attB*-Stellen flankierte DNA-Fragment. Der Rekombinationsansatz wird in *E. coli* Zellen transformiert und auf selektiven Medien ausplattiert.

Die LR-Klonierungen wurden mit dem LR ClonaseII Mix (Invitrogen) durchgeführt, von welchem 0,5  $\mu$ l in die LR-Reaktion eingesetzt wurde

#### 2.12.5 EMSA- und MST-Analysen

Die EMSA und MST-Analysen wurden von Silvia Singh (EMSA) und Christina Wolf (MST-Analysen) durchgeführt und sind beschrieben in Strauß et al., 2012.

#### 2.12.6 Isolierung von Gesamt-RNA

Für die RNA Isolierung wurde Blattmaterial geerntet, in flüssigem Stickstoff schockgefrostet und bei  $-80^{\circ}\text{C}$  aufbewahrt, sofern es nicht weiterverarbeitet wurde. Das gefrorene Blattmaterial wurde mechanisch mittels Mörser und Pistill zu einem feinen Pulver aufgeschlossen, wobei stets darauf geachtet, dass das zerkleinerte Blattmaterial nicht auftaut. Mörser und Pistill, sowie alle weiteren dafür benötigten Materialien (Löffel, Greiner-Röhrchen) waren in flüssigem Stickstoff vorgekühlt worden.

##### 2.12.6.1 Qiagen RNeasy Plant Mini Kit (für RT-PCR)

Für die Isolierung von RNA aus 50 mg Blattmaterial von *Capsicum pubescens* wurde das Qiagen RNeasy Plant Mini Kit (Qiagen) verwendet. Dabei erfolgte die Durchführung nach dem Handbuch mit folgenden Ergänzungen: Die Inkubationszeit des Pflanzenextraktes wurde, aufgrund des höheren Stärkegehalts von Paprikablättern

im Gegensatz zu Tomate, vor dem Beladen der Säule minimiert. Um DNA-Kontaminationen zu entfernen, wurde ein DNase-Verdau auf der zweiten Säule als optionaler Schritt eingefügt. Dazu wurde das Qiagen RNase-Free DNase-Set (Qiagen) nach Herstellerangaben verwendet, das speziell auf die RNeasy Kits abgestimmt ist.

#### 2.12.6.2 Trizol-Extraktion

Das durch mechanischen Aufschluss entstandene Pulver (Blattmaterial) wurde mit der entsprechenden Menge an Trizol Lösung versetzt. Dabei wurde je 100 mg Blattmaterial 1 ml Trizol zugegeben. Es wurde pro Ansatz meist 1 g Blattmaterial eingesetzt, sodass ein Volumen von 10 ml Trizol hinzugefügt wurde. Anschließend wurde der Ansatz 1 Minute kräftig gemischt und der entstandene Zellbrei für 5 min bei Raumtemperatur (RT) inkubiert. Nach der Zugabe von Chloroform (20 % des Trizol-Volumens), wurde erneut gemischt und 5 min bei RT inkubiert. Zur Trennung der Phasen wurden die Proben 15 min bei einer Geschwindigkeit von 5000 rpm bei 15 °C zentrifugiert (Sorvall Super T21, Rotor SLC- 250T). Die obere Phase wurde abgenommen, in ein neues Greiner Röhrchen überführt und zur Fällung der Nukleinsäuren mit dem gleichen Volumen an Isopropanol versetzt. Der Ansatz wurde invertiert und 10 min bei RT inkubiert. Anschließend erfolgte eine 20 minütige Zentrifugation bei einer Geschwindigkeit von 7000 rpm (Sorvall Super T21, Rotor SL-50T) und einer Temperatur von 4 °C. Der Überstand wurde abgenommen und das Pellet getrocknet. Zur Entfernung der DNA wurde ein DNase Verdau durchgeführt. Das Pellet wurde in einem Gemisch von 420 µl DEPC-Wasser, 50 µl DNase Puffer (Fermentas), 25 µl DNase (Fermentas) und 5 µl RNase-Inhibitor (Fermentas) aufgenommen und 30 min bei 37 °C inkubiert. Zum Abstoppen der Reaktion erfolgte die Zugabe von 500 µl Chloroform. Der nach Zentrifugation erhaltene Überstand wurde in ein neues Gefäß überführt und bei – 80 °C gelagert.

Trizol-Lösung:

0,8 M Guanidinthiocyanat; 0,4 M Ammoniumthiocyanat; 0,1 M Natriumacetat (pH 5); 5 % Glycerin; 38 % Phenol (in Wasser gesättigt); H<sub>2</sub>O

#### 2.12.7 Isolierung genomischer Pflanzen-DNA

Für die Isolierung genomischer DNA wurde Blattmaterial in einem Reaktionsgefäß (Eppendorf Safelock tube) schockgefrostet und anschließend mechanisch mittels Pistill, welcher vorgekühlt war, zu einem feinen Pulver aufgeschlossen. Dabei wurde darauf geachtet, dass das Reaktionsgefäß nicht beschädigt wird. Danach erfolgte die

Zugabe von 1 ml des vorgewärmten Extraktionspuffer (in Mikrowelle erhitzt) zu dem Pulver und eine extensive Vermischung mittels Vortexer. Das Reaktionsgemisch wurde im Thermomixer bei 65 °C bei 1000 rpm für mindestens 30 Minuten inkubiert und auf Eis gekühlt. Anschließend erfolgte die Zugabe von 500 µl Chloroform und eine Durchmischung durch Invertieren. Zur Phasentrennung wurde das Reaktionsgefäß 5 Minuten bei maximaler Geschwindigkeit zentrifugiert. Die obere, wässrige Phase wurde in ein neues Reaktionsgefäß, in welches 1 ml Isopropanol vorgelegt wurde, überführt und invertiert. Die DNA sollte als weiße "Fäden" sichtbar werden. Es erfolgt erneut eine Zentrifugation für 10 Minuten bei maximaler Geschwindigkeit um die DNA zu präzipitieren. Der Überstand wird verworfen und das Pellet mit 500 µl 70 %igem Ethanol gewaschen. Optional kann der Waschschrift wiederholt werden. Nach Zentrifugation (5 Minuten bei maximaler Geschwindigkeit) und Entfernen des Überstandes wird das Pellet im Thermomixer bei 65 °C getrocknet. Das Pellet (DNA) wird in 100 µl Wasser aufgenommen, im Thermomixer bei 50 °C unter leichtem Schütteln gelöst und erneut zentrifugiert. Die DNA-Lösung wird in ein neues Reaktionsgefäß überführt und wird für eine längerfristige Lagerung tiefgefroren. Die Reinheit und Menge der mit dieser Methode isolierten DNA kann anschließend mittels Agarose-Gelelektrophorese abgeschätzt werden.

#### Extraktionspuffer

0.2 M Tris/HCl pH7.5; 0.25 M NaCl; 25 mM EDTA; 0.5 % SDS (vorgewärmt auf 90 °C, Mikrowelle)

#### 2.12.8 Konzentrationsbestimmung der Nukleinsäuren

Die Konzentration und Reinheitsbestimmung von RNA und DNA erfolgte durch Absorptionsmessung bei 260 und 280 nm mittels NanoDrop ND-1000 Spektrophotometer (Thermoscientific) gegen Wasser oder DEPC-Wasser (RNA).

#### 2.12.9 cDNA Synthese für RT-PCR-Analysen

Für die cDNA-Synthese wurde der RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas) verwendet. Dabei wurde nach der Syntheseanleitung für PCR-Amplifikation verfahren. Die eingesetzte Menge an RNA betrug 500 ng und es wurden oligo dT-Oligonukleotide verwendet. Für semiquantitative RT-PCR-Analysen wurden 0,5 µl synthetisierte cDNA-Matrize (oder Verdünnungen) in eine PCR-Reaktion eingesetzt und die entsprechenden Oligonukleotide verwendet.

#### 2.12.10 RACE (*rapid amplification of cDNA ends*)

Für die Isolierung des 5'-cDNA-Endes von Bs4C-R wurde das SMARTer RACE cDNA Kit von Clontech (Takara Biotech Company) nach Herstellerprotokoll verwendet. Es wurde 1 µg Gesamt-RNA und ein genspezifischer Oligonukleotid in die cDNA-Synthese eingesetzt. Die Amplifikation der cDNA-Enden wurde mit genspezifischen Oligonukleotiden in Kombination mit adaptorspezifischen Oligonukleotiden durchgeführt (Tabelle Oligonukleotide). Die PCR-Produkte wurden in den Vektor pJET1.2 nach Herstellerprotokoll kloniert und mittels Restriktionsanalyse und Sequenzierung analysiert.

#### 2.12.11 Sequenzierung

Für die Sequenzierreaktion wurden alle Komponenten auf Eis pipettiert. Es wurden entweder 300 ng vom zu sequenzierenden Plasmid eingesetzt oder unterschiedliche Volumina von PCR-Produkten. Zu dieser Reaktion wurden 200 mM Oligonukleotid sowie Sequencing Mix (Big Dye 3.1) nach Herstellerangaben und 1,75 µl Sequencing-Puffer hinzugefügt. Das Reaktionsvolumen wurde mit Wasser auf 10 µl aufgefüllt. Die anschließende Sequenzierreaktion erfolgte in einem Thermocycler (AnalytikJena). Die Auftrennung der Fragmente, sowie die Bestimmung der Sequenz wurde vom hauseigenen Sequenzierungsservice (Sequenzierservice Ludwig-Maximilians-Universität, Biozentrum, Dr. A. Brachmann, [http://genetik.biologie.uni-muenchen.de/service/genomicsserviceunit/sequencing\\_service/index.html](http://genetik.biologie.uni-muenchen.de/service/genomicsserviceunit/sequencing_service/index.html)) durchgeführt.

PCR-Programm: 96°C/1min – [ 96°C/10sec – 50°C/15sec – 60°C/4min ] x 35 – 4°C

#### 2.12.12 Restriktionsanalysen und Ligation

Alle verwendeten Restriktionsendonukleasen und T4 DNA-Ligase wurden von Thermoscientific (Fermentas) und New England Biolabs (nur BsaI) bezogen und nach Herstellerprotokoll im entsprechenden Puffer verwendet. Die Restriktionsanalysen wurden mittels Agarose-Gelelektrophorese überprüft.

#### 2.12.13 Cut-Ligation

Alle für die cut-ligation verwendeten Plasmide wurden auf 40 fmol eingestellt und entsprechend in die Reaktion gegeben. Dem Ansatz wurde das Typ-IIs-Restriktionsenzym, T4 DNA Ligase (30 Weiss Units, Fermentas) sowie Ligase-Puffer zugegeben. Anschließend wurde das Reaktionsvolumen auf 20 µl mit Wasser

aufgefüllt. Die Reaktion erfolgte in einem Thermocycler (AnalytikJena) mit folgenden Parametern.

PCR-Programm: [ 37°C/5 Min, 20°C/5 Min ] x 50, 50°C/5 Min, 80°C/15 Min

#### 2.12.14 Erstellung der designer TALEs (dTALEs)

Die Erstellung der dTALEs erfolgte wie in (Morbiter et al., 2011) beschrieben. Die verwendeten RVDs für den dTALE AvrBs4 waren identisch mit denen des *Xcv* TALE AvrBs4. Die RVDs des dTALE Bs4C-S waren folgende:

NI-HD-HD-NI-NG-NG-NI-NS-NG-NI-NS-NG-HD-HD-NS-HD-NG-NG

Die assemblierten RVDs wurden anschließend mittels cut-ligation in einen Vektor (pENTR-D-TALE- $\Delta$ rep-BpiI-AC) integriert, der den N- und C-Terminus von AvrBs3 aber keine Repeatregion enthält. Die Erstellung der dTALEs (dTALE AvrBs4 und dTALE Bs4C-S, pENTR-D-Konstrukte) erfolgte durch Janett Elsaesser. Anschließend wurde der dTALE in den T-DNA Vektor pGWB2 mittels LR-Reaktion rekombiniert und in *A. tumefaciens* transformiert und in GUS-Analysen mit Promoter-*uidA*-Konstrukten getestet.

### 2.13 *C. annuum* cv. CM334-Genomprojekt

Die Sequenzierung des Genom des *Capsicum annuum* cv. CM334 (Criollo de Morelos 334) wurde durch die Arbeitsgruppe von Prof. Doil Choi von der Seoul Universität in Südkorea initiiert und 2014 publiziert (Kim et al., 2014). Insgesamt wurden 650,2 Gb Sequenzdaten generiert (*whole-genome shotgun*), die durch Sequenzierung von genomischen Bibliotheken mit einer durchschnittlichen Insertgröße 180 bp bis 20 kb erhalten wurde. Dies stellt eine 186-fache Abdeckung des Genoms dar. Nähere Angaben zur Erstellung der Bibliotheken und der Auswertung der Sequenzdaten sind in Kim et al. 2014 zu finden.

Die von der Firma Keygene durchgeführten RNA-Seq-Analysen mittels Illumina-Sequenzierung lieferten Sequenzdaten von potentiell AvrBs4-induzierten Gene in *C. pubescens*, die aus etwa 60 nt bestanden (siehe Anhang Read 12600). Zuvor durchgeführte RNA-Seq-Analysen, bei der 454-Sequenzierung verwendet wurde, ermöglichten größere Sequenzcontigs zu erhalten. Somit konnte einigen Illumina-Kandidatengene zusätzliche Sequenzinformationen (zu den 60 nt Reads) aus den 454-Analysen zugeordnet werden. Für die Kandidatengene der Illumina-Transkriptanalysen, für welche keine weiteren Sequenzinformationen aus den 454-

Transkriptanalysen oder der cDNA-Bibliothek erhalten werden konnte, sollte durch die Gruppe von Prof. Doil Choi homologe Sequenzen für einzelne Reads ermittelt werden. Dabei wurden mögliche Sequenzunterschiede zwischen den nah verwandten Kultivaren *C. pubescens* und *C. annuum* berücksichtigt werden. Dadurch konnte unter anderem für den Read 12600 ein homologer Bereich im Sequenzcontig 184552 identifiziert werden, der es ermöglichte Oligonukleotide abzuleiten, Amplifikationen auf genomischen Material durchzuführen und die korrespondierenden Sequenz in *C. pubescens* zu bestimmen. Zu einem späteren Zeitpunkt wurden außerdem Sequenzcontigs und deren Lage im *C. annuum*-Genom ermittelt, die weitere homologe Sequenzen zu *Bs4C-R* aufweisen (siehe 3.11).

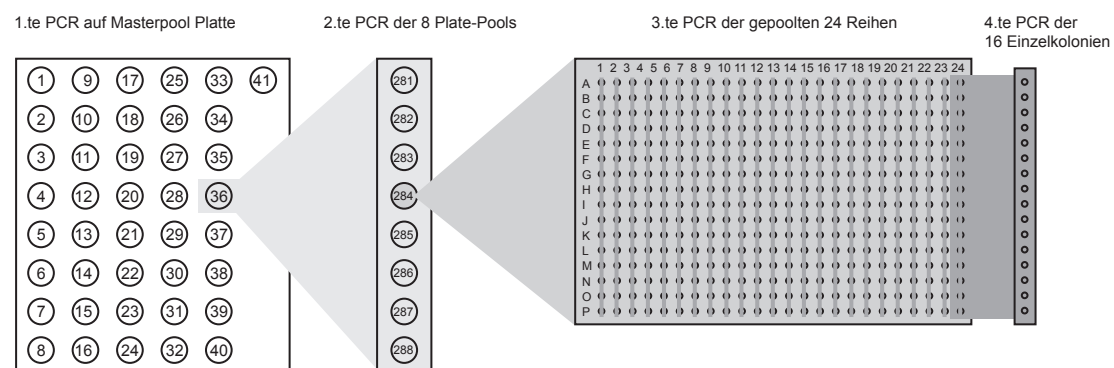
#### 2.14 Erstellung einer *C. pubescens* cDNA-Bibliothek (Invitrogen)

Für die Erstellung einer cDNA-Bibliothek wurden Blätter der *C. pubescens*-Akzession PI 235047 mit dem *avrBs4*-exprimierenden Stamm *Xcv*<sup>AvrBs4</sup> ( $5 \times 10^8$  cfu/ml) inokuliert, Blattmaterial nach 12 Stunden geerntet und RNA isoliert. Die cDNA-Synthese erfolgte mit Biotin-attB2-oligo dT und der reversen Transkriptase (SuperScript III) durch die Firma Invitrogen. Die entstandenen cDNAs wurden directional in den Vektor pDONR222.1 (Kanamycin Resistenz) kloniert. Die Bibliothek beinhaltet  $1224 \times 10^8$  cfu, wobei 91,6 % der Klone ein Insert aufwiesen. Die durchschnittliche Insertgröße der Klone beträgt 1,676 kb. Zur Analyse der cDNA-Klone kann eine BsrGI-Restriktion oder die Sequenzierung mit M13fwd+M13rev durchgeführt werden.

##### 2.14.1 Organisation der cDNA-Bibliothek

Insgesamt wurden 124416 Einzelklone der AvrBs4-induzierten *C. pubescens* cDNA-Bibliothek organisiert. Das heißt, sie wurden in 324 Mikrotiterplatten mit jeweils 384 wells im Dauermedium 2YT+HMFM (*Hogness Modified Freezing Medium*) angeordnet und bei -80°C gelagert. Die Organisation und das Spotten der cDNA-Bibliothek erfolgte in Kooperation mit Ines Walde und Patrick Schweitzer am IPK Gatersleben. Um die Isolierung von cDNA-Einzelklonen mittels PCR zu vereinfachen wurden cDNA-Klone zu verschiedenen Pools zusammengefasst. Die Superpool-Platten (4 Stück, insgesamt 324 wells) enthalten pro well die cDNA-Klone (384 Klone) von einer 384-well Plate pool-Platte. Die Masterpool-Platte enthält pro well cDNA-Klone von jeweils 8 Plate pools (3072 Klone), welche sich in einer

senkrechten Reihe befinden. Das heißt, in well 1 in der Masterpool Platte befinden sich die Plate Pools 1 bis 8. Für die Sichtung der cDNA-Bibliothek mit Oligonukleotiden für einen Kandidaten wurden zuerst 41 PCRs auf die Masterpool-Platte durchgeführt. Das positive analysierte well der Masterpool-Platte war Grundlage für die 8 PCRs auf die entsprechenden Plate-Pools der Superpool-Platte vor. Ergab die PCR ein positives Ergebnis wurde die entsprechende Plate Pool-Platte zweimal ausgestempelt und die Reihen der ersten Platte gepoolt. Mit diesen 24 Pools wurden PCR-Analysen durchgeführt. Bei Erhalten einer positiven Reihe wurden mit Hilfe der zweiten 384-Platte die Einzelklone dieser Reihe gescreent und somit konnte ein Einzelklon isoliert werden (Abb. 6).



**Abb. 6 Schema der Organisation der cDNA-Bibliothek zur Analyse mittels PCR**  
Beschreibung siehe Text. Adaptiert nach Tina Jordan 2005

## 2.15 Illumina-basierte Transkriptanalysen

Für die Transkriptanalysen wurden Blätter der *C. pubescens*-Akzessionen PI 235047 und PI 585270 mit dem WT-Xanthomonaden Stamm *Xcv* 85-10 und dem *avrBs4*-exprimierenden Stamm *Xcv*<sup>AvrBs4</sup> ( $5 \times 10^8$  cfu/mL) inokuliert und Blattmaterial nach 6, 12, 18 und 24 Stunden geerntet. Nicht-inokuliertes Blattmaterial diente als Kontrolle. Aus allen Proben wurde Gesamt-RNA mit Hilfe des RNeasy Plant Miniprep Kit (Qiagen) isoliert und daraus Transkriptbibliotheken, basierend auf der Methode beschrieben in (t Hoen et al., 2008), unter Verwendung von MseI erstellt. Für die cDNA-Synthese wurden oligo dT-Oligonukleotide eingesetzt. Die Sequenzierreaktionen wurden auf zwei Illumina GAII lanes durchgeführt wobei Proben derselben Illumina GAII lane mit spezifischen Identifikationskennzeichen versehen wurden. Die erhaltenen Sequenzen (27,6 Millionen Sequenzen/reads mit einer durchschnittlichen Länge von 76 nt aus insgesamt 18 Proben) wurden analysiert und Sequenzen mit mehr als 46 A-Nukleotiden (bspw. polyA), mit nicht auswertbaren



Nukleotidsignalen, oder mit Nukleotiden mit einem phred-Wert > 20 ausgeschlossen. Alle Sequenzen wurden auf 36 nt gekürzt und nach Entfernen des Sequenz-Identifikationskennzeichen (5 nt) in Sequenzcontigs basierend auf einer 100 % Sequenzidentität assembliert. Nur contigs mit mehr als 10 Sequenzen wurden in einer weiteren Assemblierungs-Analyse einbezogen, in denen bis zu drei Nukleotidabweichungen gebilligt werden. Dadurch sollten Sequenzierfehler und/oder Polymorphismen zwischen beiden *C. pubescens*-Akzessionen berücksichtigt werden. Nach der Assemblierungs-Analyse wurden nur contigs mit mehr als 100 Sequenzen weiterbearbeitet und Einzelproben normalisiert. Als Kandidatengene wurden nur contigs berücksichtigt, die einen 10fach höheren Transkriptlevel in der AvrBs4-induzierten *C. pubescens*-Akzession PI 235047 verglichen zur *Xcv*-WT-induzierten Akzession oder zur nicht-induzierten Kontrolle aufwiesen. Die Grundlagen und Berechnungen für die Transkriptlevel sind in Strauß et al. 2012 zu finden und die daraus erhaltenen Kandidaten im Anhang A1.

## 2.16 Verwendete Softwareprogramme

Softberry <http://linux1.softberry.com/berry.phtml>

ClustalW (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>)

Boxshade 3.21 ([http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html))

SOL Genomics Network (<http://www.sgn.cornell.edu/tools/blast/>)

National Center for Biotechnology Information (NCBI) ([www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/))

Expasy Translate Tool (<http://web.expasy.org/translate/>)

Sequencher und CLC workbench

Sequence Massager (<http://www.attotron.com/cybertory/analysis/seqMassager.htm>)

t-coffee (<http://tcoffee.crg.cat/apps/tcoffee/index.html>)

Transmembran-Helices-Vorhersageprogramm nach dem Hidden-Markov-Model  
<http://www.cbs.dtu.dk/services/TMHMM-2.0/>

Vorhersageprogramme für EBE für TAL-Effektoren:

Talgetter, <http://galaxy.informatik.uni-halle.de> (Grau et al., 2013)

Storyteller, <http://bioinfo-prod.mpl.ird.fr/xantho/tales>, (Perez-Quintero et al., 2013)

Target Finder TALE-NT 2.0 suite, <https://tale-nt.cac.cornell.edu/>, (Doyle et al., 2012)

Talvez, <http://bioinfo.mpl.ird.fr/cgi-bin/talvez/talvez.cgi>, (Perez-Quintero et al., 2013)

### 3 Ergebnisse

#### Teile der vorliegenden Arbeit wurden bereits publiziert in:

Strauß T, Van Poecke R, Strauß A, Römer P, Minsavage GV, Singh S, Wolf C, Strauß A, Kim S, Lee H-A, Yeom S-I, Parniske M, Stall RE, Jones JB, Choi D, Prins M, Lahaye T.

RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large crop genome. Proc Natl Acad Sci. 2012 Oct; 109: 19480 - 19485

Die Durchführung und Auswertung der Experimente wurden, mit Ausnahme der Transkriptanalysen, der TALE-DNA-Bindestudien (EMSA- und MST-Analysen) und der Erstellung des dTALE *EBE Bs4C-S* zur Induktion des *EBE Bs4C-S*, von mir durchgeführt. Die Transkriptanalysen und deren Auswertung wurden durch die Firma Keygene von Remco van Poecke und Marcel Prins durchgeführt. Das dafür benötigte Material (RNA) wurde von mir bereitgestellt. Die EMSA-Analysen wurden von Sylvia Singh aus der Arbeitsgruppe von Prof. Martin Parniske und Christina Wolf aus der Arbeitsgruppe von Prof. Thomas Lahaye durchgeführt (siehe 3.7, Abb. 15, siehe 3.13.1). Die MST-Analysen (siehe 3.7, Abb. 15) wurden von Christina Wolf aus der Arbeitsgruppe von Prof. Thomas Lahaye durchgeführt. Die Erstellung des dTALE *EBE Bs4C-S* zur Induktion des *EBE Bs4C-S* (siehe 3.9) erfolgte durch Janett Elsaesser aus der Arbeitsgruppe von Prof. Thomas Lahaye.

Übersicht der modifizierten Tabellen und Abbildungen in der Dissertation, welche bereits in Strauß et al., 2012 publiziert wurden.	
<b>Tabelle/Abbildung/Anhang in der Dissertation von Tina Strauß</b>	<b>Tabelle/Abbildung/Anhang (supplement) in der Publikation Strauß et al., 2012</b>
Tab. 1	Tab. S2 modifiziert
Abb. 5	Fig. S4 modifiziert
Abb. 7	Fig. 1 modifiziert
Abb. 9	Fig. S7 modifiziert
Abb. 12	Fig. S8 modifiziert
Abb. 13	Fig. 3 modifiziert
Abb. 14	Fig. S9 modifiziert
Abb. 15	Fig. S10 modifiziert
Anhang A1	Tab. S2 modifiziert
Anhang A3	Tab. S1 modifiziert

Anhang A4	SI Text 1 modifiziert
Anhang A5	SI Text 3 modifiziert
Anhang A6	SI Text 3 und SI Text 5 modifiziert
Anhang A34	Tab. S3 modifiziert
Modifiziert bedeutet, dass entweder die Sprache, die Anordnung, die Abbildung oder die Tabelle selbst, wenn auch nur minimal, auf Basis der publizierten Abbildungen/Tabellen verändert wurden.	

### 3.1 Isolierung AvrBs4-induzierter Genen mittels Transkriptomprofiling

Der *Xcv* TAL-Effektor AvrBs4 induziert in der *C. pubescens*-Akzession PI 235047 eine HR (Minsavage et al, 1999; Stall et al, 2009). Dabei konnte gezeigt werden, dass diese HR abhängig von funktionalen NLS und AD ist (siehe 1.8), was nahe legt, dass AvrBs4 die Expression seines korrespondierenden *R*-Gens, bezeichnet als *Bs4C-R*, induziert (Gürlebeck, 2001). Eine transkriptionelle Induktion wurde schon für die *R*-Gene *Bs3* aus *C. annuum* und *Xa27* aus *O. sativa* gezeigt. Dabei waren die *R*-Gen Transkripte ausschließlich in Anwesenheit des jeweiligen TAL-Effektors nachweisbar, während Transkripte von TAL-induzierten *S*-Genen auch in Abwesenheit des TAL-Effektors detektierbar waren (Gu et al., 2005; Kay et al., 2007; Römer et al., 2007). Dies weist darauf hin, dass die Transkription dieser *R*-Gene "streng" reguliert ist, was in Anbetracht der Tatsache, dass die korrespondierenden *R*-Proteine *Bs3* und *Xa27* Zelltod induzieren, notwendig ist. Basierend auf der Annahme, dass *Bs4C-R* ebenfalls in seiner Expression einer strengen Regulation unterliegt, d.h. nur in Anwesenheit von AvrBs4 induziert wird, wurden Illumina basierte Transkriptomanalysen initiiert, mit dem Ziel *Bs4C-R* als AvrBs4-induziertes Gen zu isolieren.

Für die Transkriptomanalysen wurden die *C. pubescens*-Akzessionen PI 235047 (AvrBs4-resistent) und PI 585270 (AvrBs4-suszeptibel) mit *Xcv* 85-10 Wildtyp (WT) und *avrBs4*-exprimierenden Xanthomonaden inokuliert und Blattmaterial nach 6, 12, 18 und 24 Stunden geerntet. Nicht inokuliertes Blattmaterial diente als Kontrolle. Für alle 18 RNA-Proben wurden Transkriptbibliotheken durch unseren Kooperationspartner Keygene von Remco van Poecke und Marcel Prins erstellt (siehe 2.15) (Strauß et al., 2012; t Hoen et al., 2008). Die cDNA-Synthese erfolgte dabei ausgehend von oligo dT-Oligonukleotiden. Die Sequenzierungen aller 18 Proben ergab insgesamt 27,6 Millionen Sequenzen/reads mit einer durchschnittlichen Länge von 76 Nukleotiden (nt) (siehe 2.15). Sequenzen geringer Qualität wurden nicht in die weiteren Auswertungen mit einbezogen. Die verbliebenen Sequenzen wurden nach Entfernen des Sequenz-Identifikationskennzeichen anhand der Übereinstimmung der ersten 31 nt in contigs assembliert. Dabei wurden bis zu drei Nukleotidabweichungen zwischen den Sequenzen toleriert, die aufgrund möglicher Sequenzierfehler und/oder

Unterschiede zwischen beiden *C. pubescens*-Akzessionen auftreten können (Strauß et al., 2012). Da die Transkription von *Bs4C* durch AvrBs4 induziert wird und demnach der Transkriptlevel stark ansteigen sollte, wurden im Folgenden nur Contigs analysiert, die mehr als 100 Sequenzen aufwiesen – das waren 16220. Von diesen zeigten 32 Contigs eine zehnfache Transkriptinduktion in der resistenten Akzession gegenüber der suszeptiblen Akzession bzw. der Expressionskontrolle (siehe A1, Strauß et al., 2012). Jedoch erfüllte nur einer dieser 32 potentiellen *Bs4C-R*-Kandidaten unsere Kriterien für ein AvrBs4-induziertes *R*-Gen (Tab. 1). Kandidat 12600 war nur nach Translokation von AvrBs4 exprimiert und ausschließlich in der resistenten Akzession PI 235047. Transkripte konnten erstmals 18 Stunden nach Inokulation der *Xanthomonas*-Bakterien nachgewiesen werden (Tab. 1).

Pflanzen-Genotyp	<i>C. pubescens</i> PI 235047										<i>C. pubescens</i> PI 585270										Max C <sup>2</sup>	PI 235047 Xcv <sup>AvrBs4</sup> 24 hpi / Max Xcv				
	Inokulation mit	C					Xcv					Xcv <sup>AvrBs4</sup>					C	Xcv					Xcv <sup>AvrBs4</sup>			
Cluster	hpi	0	6	12	18	24	6	12	18	24	0	6	12	18	24	6	12	18	24	0	6	12	18	24		
1.	12600	0.0	0.0	0.0	0.8	0.0	0.0	0.0	33.6	87.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		

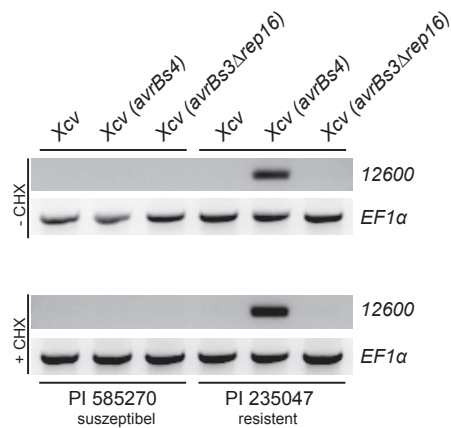
**Tab. 1 Das Expressionsmuster des *Bs4C*-Kandidatengens 12600**  
Die Transkriptanzahl für den *Bs4C*-Kandidaten 12600 ist in Abhängigkeit vom Genotyp der Pflanze (*C. pubescens* PI 235047 oder PI 585270), von der Behandlung (*Xcv* oder *Xcv*<sup>AvrBs4</sup>) und von den verschiedenen Erntezeitpunkten (6,12,18 und 24 hpi) dargestellt. hpi - Stunden nach Inokulation (*hours post inoculation*); C – Kontrolle (nicht inokuliert), *Xcv* – *Xanthomonas campestris* pv. *vesicatoria*  
Abbildung modifiziert nach Strauß et al., 2012 Tab.S2.

3.2 RT-PCR-Analysen zur Bestätigung des Kandidatengens 12600 als AvrBs4-induziertes Gen

Das Cluster 12600 umfasste nur 62 Nukleotide (siehe A4). In Zusammenarbeit mit der Arbeitsgruppe von Prof. Doil Choi, die an der Sequenzierung und Assemblierung des *Capsicum annuum* cv. CM334 Genoms arbeitet, konnte jedoch ein größerer Sequenzcontig (contig 184552, siehe A4) identifiziert werden, der einen homologen Bereich zu dem *C. pubescens* cluster 12600 aufwies. Basierend auf dieser Sequenz wurden Oligonukleotide (siehe A4, A34) für den *Bs4C*-Kandidaten 12600 abgeleitet und semiquantitative Reverse Transkriptase-PCR (sq RT-PCR)-Analysen mit nicht-induziertem und AvrBs4-induziertem Material der resistenten und suszeptiblen *C. pubescens*-Akzessionen durchgeführt werden. Zur Vereinfachung wird der *Bs4C*-Kandidat 12600 der AvrBs4-resistenten Akzession (PI 235047) als *Bs4C-R* und der aus der AvrBs4-suszeptiblen Akzession (PI 585270) als *Bs4C-S* bezeichnet. Die RT-PCR-Analysen zeigten, dass der *Bs4C-R*-Kandidat 12600 in der Tat in der resistenten Akzession und nur durch AvrBs4 induziert wird (Abb. 7). In der suszeptiblen

Akzession konnte weder in Abwesenheit noch in Anwesenheit von AvrBs4-Transkript nachgewiesen werden (Abb. 7). Die Abwehrreaktion der Pflanze geht mit einer Umprogrammierung der gesamten Transkriptmaschinerie einher. Um zu zeigen, dass die Induktion des *Bs4C-R*-Kandidaten 12600 spezifisch für AvrBs4 ist und nicht durch die pflanzliche Abwehrreaktion hervorgerufen wird, wurde AvrBs3Δrep16-induziertes Material auf Präsenz des *Bs4C-R*-Kandidatentranskript mittels sq RT-PCR getestet. Für AvrBs3Δrep16 konnte gezeigt werden, dass es in beiden *C. pubescens*-Akzessionen eine HR durch Bindung und Aktivierung des *Bs3-E*-Promoter induziert (Römer, 2010; Strauß et al., 2012). Wie in Abb. 7 erkennbar, ist die Induktion des *Bs4C-R*-Kandidatengens 12600 spezifisch nur durch AvrBs4 induziert und nicht durch die AvrBs3Δrep16-induzierte Resistenzreaktion (Strauß et al., 2012).

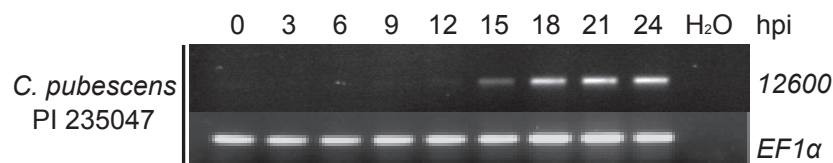
Die sq RT-PCR-Analysen zeigten, dass *Bs4C-R*-Transkript des Kandidatengens 12600 in der resistenten Akzession durch AvrBs4 induziert wird. Eine Induktion von *Bs4C-R* könnte sowohl direkt, durch Bindung und Aktivierung von AvrBs4 an den *Bs4C-R*-Promoters, als auch indirekt, durch Proteine, die von primären AvrBs4-Zielgenen kodiert werden, erfolgen (siehe 4.2.3, Abb. 26). Um dies zu prüfen, wurden sq RT-PCR-Analysen nach Koinfiltration von Cycloheximid (CHX), welches die Protein-Neusynthese und somit die Induktion von sekundären Zielgenen inhibiert, durchgeführt. Dies ermöglicht die Unterscheidung zwischen direkten und indirekten Zielgenen (siehe 4.2.3, Abb. 26). Die sq RT-PCR-Studien (Abb. 7) zeigen, dass Transkript des *Bs4C-R*-Kandidatengens 12600 in der resistenten Akzession auch bei CHX-Infiltration nachgewiesen werden konnte. Dies deutet darauf hin, dass das *Bs4C-R*-Kandidatengens 12600 ein direktes Zielgen von AvrBs4 ist und ein EBE für AvrBs4 im Promoter aufweisen sollte.



**Abb. 7 Bestätigung des *Bs4C-R*-Kandidaten 12600 als direktes *AvrBs4*-Zielgen mittels sq RT-PCR-Analysen**

*C. pubescens* Blätter wurden mit *Xcv*-Stämmen, die verschiedene TALE-Gene exprimieren (*AvrBs4* und *AvrBs3Δrep16*), mit und ohne Cycloheximid (+/- CHX), infiziert und 18 Stunden nach Inokulation geerntet. Aus diesem Material wurde cDNA synthetisiert, welches die Grundlage für die sq RT-PCR Analysen war. Für die Durchführung der sq RT-PCR Analysen wurden Oligonukleotide zur Amplifikation eines Fragments des Illumina clusters 12600 (*Bs4C*) verwendet (siehe A4, A34). Als Kontrolle für den Einsatz vergleichbarer cDNA-Mengen wurden Oligonukleotide für den Elongationsfaktor 1α (*EF1α*) verwendet (A34). Die erhaltenen PCR-Produkte wurden auf einem 1 % igen Agarosegel aufgetrennt und mit Ethidiumbromid angefärbt. Abbildung modifiziert nach Strauß et al., 2012 Fig. 1.

Mittels sq RT-PCR Studien wurde außerdem untersucht, wann Transkript des *Bs4C-R* Kandidaten 12600 erstmals detektierbar ist. Dafür erfolgte die Ernte von induziertem Blattmaterial der resistenten *C. pubescens*-Akzession alle drei Stunden bis 24 Stunden nach Infiltration mit *avrBs4*-exprimierenden Xanthomonaden. Die Daten des Transkriptprofiling ließen vermuten, dass der Zeitpunkt zwischen 12 und 18 Stunden liegen sollte (Tab. 1). Wie Abb. 8 zeigt, konnte Transkript des *Bs4C-R*-Kandidaten 12600 15 Stunden nach *AvrBs4*-Induktion nachgewiesen werden.



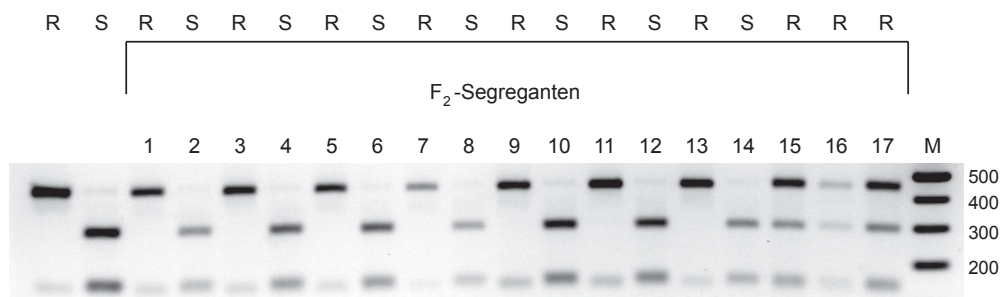
**Abb. 8 Transkriptnachweis des *Bs4C*-Kandidatengens 15 Stunden nach *AvrBs4* Induktion**

*C. pubescens* Blätter der resistenten Akzession (PI 235047) wurden mit *avrBs4*-exprimierenden Xanthomonaden infiziert und 3, 6, 9, 12, 15, 18, 21 und 24 Stunden nach Inokulation (hpi) geerntet. Für die Durchführung der sq RT-PCR Analysen wurden Oligonukleotide zur Amplifikation des *Bs4C* Kandidatengens 12600 verwendet (siehe A4, A34). Als Kontrolle für den Einsatz vergleichbarer cDNA Mengen wurden Oligonukleotide für den Elongationsfaktor 1α (*EF1α*) verwendet (A34). Die erhaltenen PCR-Produkte wurden auf einem 1 %igen Agarosegel aufgetrennt und mit Ethidiumbromid angefärbt.



### 3.3 Die genetische Kartierung des Kandidaten 12600 bestätigt die Kopplung mit der AvrBs4-induzierten Resistenzreaktion

Im Folgenden wurden mit den RT-PCR-Oligonukleotiden genomische Fragmente für beide *C. pubescens*-Akzessionen amplifiziert und vergleichend sequenziert (siehe A5, A6). Ziel war es, mögliche Polymorphismen zwischen beiden Akzessionen zu identifizieren und für die genetische Kartierung des *Bs4C-R*-Kandidaten 12600 zu verwenden. Das amplifizierte Fragment umfasste etwa 450 Nukleotide und wies sieben Nukleotidpolymorphismen und acht Insertions/Deletionspolymorphismen (InDel) auf (siehe A6). Einer der Nukleotidpolymorphismen resultiert in einem *Bgl*III-Restriktionsfragmentlängenpolymorphismus (RFLP) zwischen der suszeptiblen und der resistenten Akzession (siehe A5). Dieser wurde zur Etablierung eines CAPS-Markers genutzt. Der CAPS-Marker wurde für die Analyse von 24 Pflanzen einer F<sub>2</sub>-Population, die aus einer Kreuzung von suszeptiblem und resistentem Elter hervorgegangen ist, verwendet (Abb. 9, siehe A3) (Gürlebeck, 2001). Wie in Abb. 9 erkennbar, zeigte der CAPS-Marker bei allen phänotypisch-suszeptiblen getesteten F<sub>2</sub>-Segreganten (S) das Restriktionsmuster des suszeptiblen Elters (S, geschnitten, 2 Fragmente). Bei allen resistenten F<sub>2</sub>-Pflanzen (R) hingegen wurde entweder die Allelkonfiguration des resistenten Elters (R, homozygote Nachkommen, nicht geschnitten, 1 Fragment) oder beider Eltern (heterozygote Nachkommen, 3 Fragmente) erhalten. Diese Daten lassen vermuten, dass der *Bs4C-R*-Kandidat genetisch gekoppelt ist mit der AvrBs4-induzierten HR und es nahe liegt, dass der Kandidat 12600 das Resistenzgen *Bs4C-R* darstellt.



**Abb. 9 Der *Bs4C-R*-Kandidat ist genetisch gekoppelt mit der AvrBs4-induzierten Immunantwort**  
Im 3' Bereich des *Bs4C*-Kandidatengens konnte ein CAPS (*cleaved amplified polymorphic sequence*)-Marker (Restriktionspolymorphismus-*Bgl*III) identifiziert werden. Mit diesem CAPS-Marker wurden die *C. pubescens*-Eltern und 24 F<sub>2</sub>-Pflanzen einer Kreuzungspopulation von den *C. pubescens*-Akzession PI 235047 (R) und PI 585270 (S), getestet. Auf genomische DNA der Eltern-Pflanzen und der 24 F<sub>2</sub>-Segreganten erfolgte die Amplifikation eines Fragmentes mit spezifischen Oligonukleotiden (siehe A5). Das resultierende Fragment schloss den *Bgl*III-Polymorphismus mit ein, sodass anschließend eine *Bgl*III-Restriktionsanalyse durchgeführt werden konnte. Die erhaltenen Restriktionsfragmente wurden auf einem 3 %igen Agarosegel aufgetrennt und zur Visualisierung mit

Ethidiumbromid gefärbt. Für 17 Segreganten wurde das Restriktionsmuster hier exemplarisch dargestellt. Die vollständige Analyse ist im Anhang A3 zu finden.

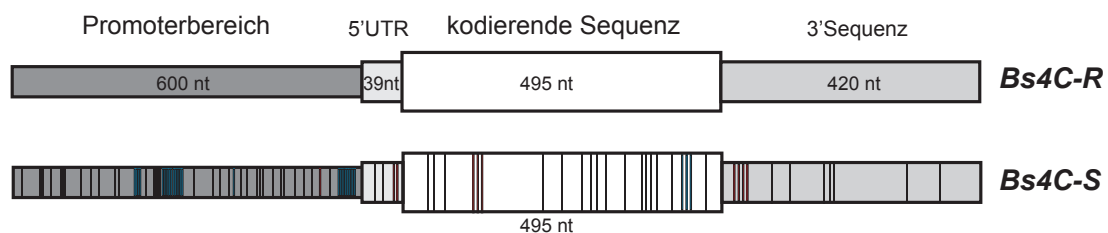
M = Größenmarker (GeneRuler 50 bp DNA Größenstandard-Fermentas)

Abbildung modifiziert nach Strauß et al., 2012 Fig.S7.

### 3.4 Das *Bs4C*-Kandidatengen 12600 vermittelt die Erkennung von AvrBs4

#### 3.4.1 Isolierung und vergleichende Sequenzierung des *Bs4C*-Kandidatengen aus der suszeptiblen und resistenten *C. pubescens*-Akzession

Zur funktionellen Charakterisierung des *Bs4C*-Kandidaten war es notwendig seine Genstruktur zu bestimmen. Mit Hilfe des vorliegenden *C. annuum*-Sequenzcontig 184552 (siehe A4) konnte die kodierende Sequenz (KDS) des *Bs4C-R*-Homologs *CaBs4C.1* vorhergesagt werden (Softberry Programm, <http://linux1.softberry.com/berry.phtml>), die unter anderem Grundlage für die Bestimmung der *C. pubescens* *Bs4C-R*-Sequenz war. Parallel wurden RACE-Analysen (*rapid amplification of cDNA ends*) zur Bestimmung des Transkriptionsstarts und Genome Walker-Experimente zur Ermittlung der Promotersequenz durchgeführt (siehe A5). Unter Verwendung dieser Ansätze konnten etwa 1500 nt-Sequenz für *Bs4C-R* erhalten werden, die 600 nt Promoter, 39 nt 5'UTR, eine KDS von 495 nt und eine 3'Sequenz von 420 nt umfasst. Ein Vergleich der gDNA mit cDNA machte deutlich, dass *Bs4C-R* lediglich aus einem Exon besteht und keine Introns aufweist (siehe A5+A6).



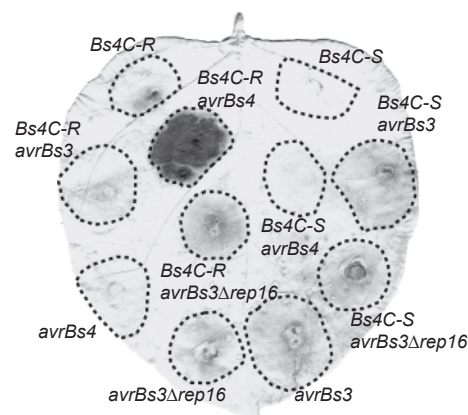
**Abb. 10 Genstruktur von *Bs4C-R* vergleichend zu *Bs4C-S***

Aufgrund vergleichender Sequenzierungen, RACE-Analysen und Genome Walker-Experimenten konnte die Sequenz, der Transkriptionsstart und die Genstruktur von *Bs4C-R* bestimmt werden. Die dargestellten Boxen stellen unterschiedliche Elemente dar, den Promoterbereich, die 5'UTR (*untranslated region*), die KDS und die 3'Sequenz. Senkrechte schwarze Striche markieren Nukleotid-, rote Striche Insertions- und blaue Striche Deletionspolymorphismen, die *Bs4C-S* verglichen zu *Bs4C-R* aufweist. Für *Bs4C-S* kann keine genaue Angabe über die Größe der 5'UTR gemacht werden.

Der Vergleich der Nukleotidsequenzen der *Bs4C*-Allele aus der resistenten (*Bs4C-R*) und der suszeptiblen (*Bs4C-S*) *C. pubescens*-Akzession zeigt, dass Polymorphismen sowohl in den Promoter-, den kodierenden als auch den UTR-Sequenzen auftreten (Abb. 10, siehe A6).

### 3.4.2 Komplementationanalysen bestätigen den Kandidaten 12600 als Resistenzgen *Bs4C-R*

Für die Komplementationsanalysen wurden Promoter-KDS-Fragmente beider Akzessionen (280 nt Promoter + 495 nt KDS mit Stopp) von gDNA amplifiziert, in den T-DNA Vektor pGWB3 transferiert und anschließend in den *A. tumefaciens* Stamm GV3101 transformiert. Beide *Bs4C*-Allele wurden anschließend allein oder zusammen mit *35S:avrBs4*, *35S:avrBs3* und *35S:avrBs3Δrep16* *Agrobacterium*-vermittelt transient in *N. benthamiana* exprimiert. Wie Abb. 11 zeigt, resultiert ausschließlich die Ko-Transformation von *Bs4C-R* mit *35S:avrBs4* in einer HR. Weder die Expression von *Bs4C-R* allein, noch die Ko-Transformation von anderen TALE-Genen mit *Bs4C-R*, führt zu einer HR. Das *Bs4C-S*-Allel aus der suszeptiblen Akzession induziert keine HR - weder alleine noch in Ko-Expression mit *avrBs4* (Abb. 11). Zusammenfassend zeigen die Komplementationsanalysen, dass nur *Bs4C-R* und nicht *Bs4C-S* in Kombination mit *avrBs4* eine HR auslösen kann, was bedeutet, dass *Bs4C-R* tatsächlich das Resistenzgen aus der *C. pubescens*-Akzession PI 235047 darstellt.



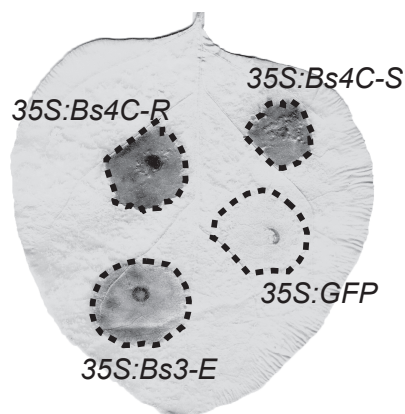
**Abb. 11 Das *Bs4C-R*-Kandidatengene vermittelt die Erkennung von *AvrBs4*.**

Phänotypen nach *Agrobacterium*-vermittelter Transformation von *N. benthamiana*. Die in einen T-DNA Vektor klonierten *Bs4C*-Kandidatengene (Promoter und KDS) aus der resistenten (*Bs4C-R*) und der suszeptiblen (*Bs4C-S*) *C. pubescens*-Akzession wurden entweder allein oder in Kombination mit verschiedenen TALE-Genen (*avrBs3*, *avrBs3Δrep16* und *avrBs4*) transfiziert. Die TALE-Gene standen unter Kontrolle des konstitutiven *Cauliflower Mosaic Virus (CMV)* 35S-Promoters. Inokulierte Blattbereiche wurden mit einer gestrichelten Linie gekennzeichnet. Die Blätter wurden vier Tage nach Inokulation geerntet und zum besseren Sichtbarmachen der HR (dunkle Blattbereiche) in Ethanol gebleicht.

### 3.5 Die Überexpression der *Bs4C*-Allele resultiert in einer HR

Um zu prüfen, ob Polymorphismen zwischen *Bs4C-R* und *Bs4C-S* in den Promotoren oder der KDS oder in beiden für die differentielle *AvrBs4*-Erkennung verantwortlich

ist (Abb. 10, Abb. 11), wurden die zwei KDS unter transkriptioneller Kontrolle des konstitutiven *Cauliflower Mosaic Virus (CMV)* 35S-Promoters *Agrobacterium*-vermittelt transient in *N. benthamiana* exprimiert (siehe A6). Als Kontrollen diente *Bs3-E* (35S:*Bs3-E*), welches bei Überexpression eine HR induziert (Abb. 12) und ein 35S:*GFP*-Konstrukt (*green fluorescent protein*) als Negativkontrolle (Römer et al., 2007). Sowohl *Bs4C-R* als auch *Bs4C-S* lösen bei konstitutiver Expression eine HR aus, was darauf hinweist, dass Polymorphismen im Promoterbereich und nicht in der KDS für die differentielle AvrBs4-Erkennung verantwortlich sind (Abb. 12).



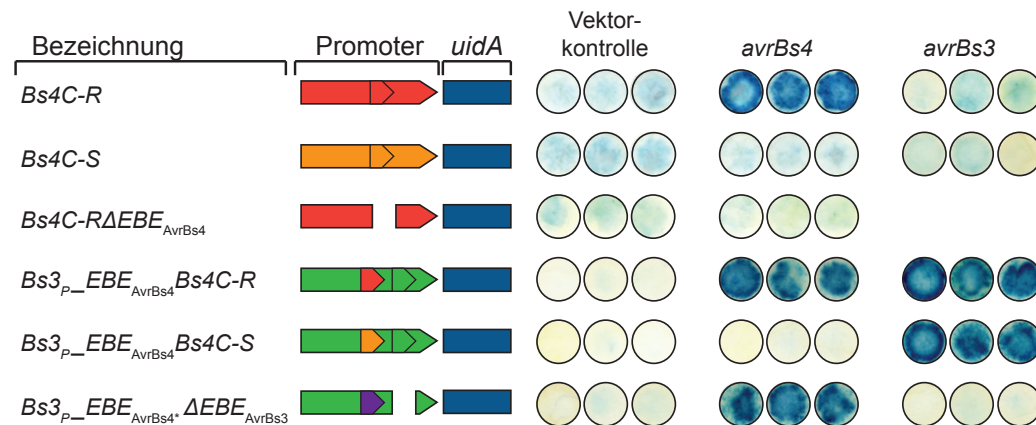
**Abb. 12 Die konstitutive Expression der *Bs4C*-Allele resultiert in der Ausbildung einer HR**

Die KDS des resistenten (*Bs4C-R*) und des susceptible (*Bs4C-S*) *C. pubescens*-Allels wurden in einen T-DNA Vektor unter Kontrolle des 35S-Promoters kloniert. Als Negativkontrolle wurde ein konstitutiv exprimiertes GFP verwendet. Die konstitutive Expression eines Resistenzgens aus dem *C. annuum*-Kultivar ECW (*Bs3-E*) diente als Positivkontrolle (Römer et al., 2007). Inokulierte Blattbereiche wurden mit einer gestrichelten Linie gekennzeichnet. Die Blätter wurden drei Tage nach Inokulation geerntet und zum besseren Sichtbarmachen der HR (dunkle Blattbereiche) in Ethanol gebleicht.

Abbildung modifiziert nach Strauß et al., 2012 Fig.S8.

Im Folgenden wurden 280 bp Promoter-Fragmente (5' vom ATG; siehe A5) beider Allele amplifiziert und *upstream* des Reportergens *uidA* in den T-DNA-Vektor pGWB3 kloniert. *uidA* kodiert für eine  $\beta$ -Glucuronidase und seine Expression kann mittels X-Gluc-Substratumsatz detektiert werden. Bindet und aktiviert AvrBs4 einen oder beide Promotoren, wird dies durch eine Blaufärbung der infiltrierten Blattbereiche sichtbar. Die *Bs4C-R*- und *Bs4C-S*-Promoter-*uidA*-Konstrukte wurden transient in *N. benthamiana* durch *Agrobacterium*-vermittelte Transformation allein oder in Kombination mit den TALE-Genen *avrBs3* und *avrBs4* exprimiert. Wie in Abb. 13 an den blau gefärbten Blattscheiben zu erkennen ist, wird nur der *Bs4C-R*-Promoter durch AvrBs4 aktiviert, jedoch nicht der *Bs4C-S*-Promoter. Die Aktivierung von *Bs4C-R* ist AvrBs4-spezifisch, da eine Ko-Expression mit dem TALE-Gen *avrBs3* nicht in einer Blaufärbung resultiert (Abb. 13).

Die GUS-Analysen bestätigen damit, dass die funktional relevanten Polymorphismen zwischen beiden *Bs4C*-Allelen im Promoter und nicht in der KDS liegen. Außerdem zeigt die AvrBs4-abhängige transkriptionelle Aktivierung des *Bs4C-R*-, jedoch nicht des *Bs4C-S*-Promoters, dass allein der *Bs4C-R*-Promoter die “Erkennung” von AvrBs4 vermittelt.



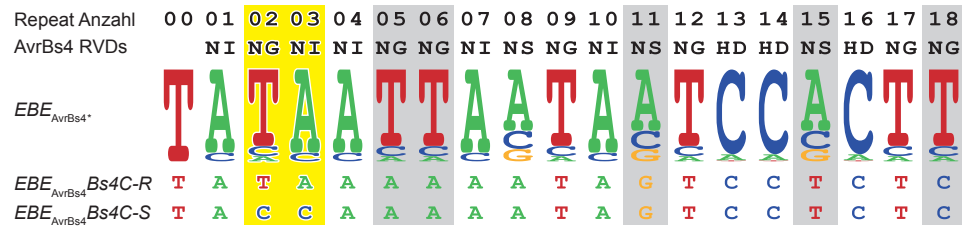
**Abb. 13 *Bs4C-R* enthält ein Effektorbindeelement (EBE), welches notwendig und ausreichend für die transkriptionelle Aktivierung durch AvrBs4 ist.**

GUS-Analysen nach *Agrobacterium*-vermittelter Transformation von *N. benthamiana* von EBE-Promoter:uidA-Reporter-Konstrukten und TALE-Genen. Die graphische Darstellung der analysierten EBE-Promoter:uidA-Reporter-Konstrukte ist auf der linken Seite zu finden. Rot und orange markiert sind die Promotoren von *Bs4C-R* und *Bs4C-S*. Pfeile innerhalb der Promotoren stellen das EBE von AvrBs4 bzw. die korrespondierende Sequenz im *Bs4C-S*-Promoter dar. Das EBE<sub>AvrBs4\*</sub> bezeichnet das für AvrBs4 anhand des TALE-Codes vorhergesagte Bindeelement (lila Pfeil). Der *Bs3*-Promoter ist grün markiert und der Pfeil innerhalb zeigt das EBE für AvrBs3. In den Konstrukten *Bs4C-R*ΔEBE<sub>AvrBs4</sub> und *Bs3*<sub>P</sub>-EBE<sub>AvrBs4\*</sub>ΔEBE<sub>AvrBs3</sub> wurden die EBE deletiert, währenddessen sie in den Konstrukten *Bs3*<sub>P</sub>-EBE<sub>AvrBs4\*</sub>ΔEBE<sub>AvrBs3</sub>, *Bs3*<sub>P</sub>-EBE<sub>AvrBs4</sub>*Bs4C-R* und *Bs3*<sub>P</sub>-EBE<sub>AvrBs4</sub>*Bs4C-S* inseriert wurden. Die EBE-Promoter-uidA T-DNA-Konstrukte wurden in Kombination mit leerem Vektor, den TALE-Genen *avrBs4* oder *avrBs3* unter Kontrolle des 35S-Promoter in der Pflanze exprimiert. Blattscheiben wurden 38 Stunden nach Infiltration geerntet und mit X-Gluc (Cyclohexylammoniumsalz der 5-Brom-4-chlor-3-indolyl-β-D-glucuronsäure) gefärbt und in Ethanol gebleicht um die Blaufärbung hervorzuheben.

Abbildung modifiziert nach Strauß et al., 2012 Fig.3.

### 3.6 Das EBE<sub>AvrBs4</sub>*Bs4C-R* vermittelt die AvrBs4-Erkennung

Der Vergleich der Promotoren beider Allele (*Bs4C-R* und *Bs4C-S*) zeigte eine große Anzahl von Nukleotid- sowie InDel-polymorphismen (Abb. 10, siehe A6). Mit Hilfe des TALE-Codes wurde der *Bs4C-R*-Promoter genauer analysiert und es konnte in der Promotersequenz ein potentiell EBE für AvrBs4 (EBE<sub>AvrBs4</sub>*Bs4C-R*) 123 Nukleotide 5' vom Transkriptionsstart identifiziert werden (Abb. 14, siehe A5). Das 19 Nukleotide umfassende EBE<sub>AvrBs4</sub>*Bs4C-R* weist im Vergleich mit einem anhand des TALE-Code vorhergesagtem AvrBs4-EBE (EBE<sub>AvrBs4\*</sub>) fünf Polymorphismen auf, die in der Abb. 14 grau unterlegt sind.



**Abb. 14 Die Nukleotidsequenz des  $EBE_{AvrBs4}Bs4C-R$  weist fünf Polymorphismen zu dem anhand des TALE-Codes vorhergesagten  $EBE_{AvrBs4^*}$  auf**

Die Abbildung zeigt die RVDs von AvrBs4 (oben) und das anhand des TALE-Codes (siehe 1.4, Abb. 2) abgeleitete „perfekte“  $EBE$  für AvrBs4 ( $EBE_{AvrBs4^*}$ ). Darunter sind die Nukleotidsequenzen des im  $Bs4C-R$ -Allel identifizierten  $EBE_{AvrBs4}$  und die korrespondierende Sequenz im  $Bs4C-S$ -Allel dargestellt. Grau hinterlegt sind Nukleotide, die von dem „perfekten“  $EBE_{AvrBs4^*}$  abweichen. Die Unterschiede zwischen dem  $EBE_{AvrBs4}Bs4C-R$  und  $EBE_{AvrBs4}Bs4C-S$  an Position 2 und 3 sind gelb unterlegt.

Abbildung modifiziert nach Strauß et al., 2012 Fig.S9.

Um zu hinterfragen, ob das identifizierte potentielle  $EBE_{AvrBs4}Bs4C-R$  das funktionale Element im  $Bs4C-R$ -Promoter ist, wurde ein  $Bs4C-R$ -Promoter:*uidA*-Konstrukt erstellt, in dem das  $EBE_{AvrBs4}Bs4C-R$  deletiert wurde. GUS-Analysen belegten, dass das Deletionskonstrukt ( $Bs4C-R\Delta EBE_{AvrBs4}$ ) nicht mehr durch AvrBs4 aktivierbar wird (Abb. 13).

Um weiterhin zu demonstrieren, dass das  $EBE_{AvrBs4}Bs4C-R$  nicht nur notwendig sondern auch ausreichend für die AvrBs4-vermittelte Promoteraktivierung ist, wurde es in einen anderen Promoterkontext - in den Promoter des Resistenzgens  $Bs3$  aus Paprika integriert (Römer et al., 2009a). Für den  $Bs3$ -Promoter konnte in früheren Experimenten gezeigt werden, dass dieser transkriptionell durch AvrBs3, einem *Xcv*-TALE mit einer Identität von 97 % zu AvrBs4, nicht aber durch AvrBs4, aktiviert wird (Römer et al., 2007).

Das erstellte  $Bs3$ -Promoterinsertionskonstrukt ( $Bs3_P\_EBE_{AvrBs4}Bs4C-R$ ) wurde in transienten GUS-Analysen allein oder mit AvrBs3 bzw. mit AvrBs4 getestet (Abb. 13). Im Vergleich zum WT- $Bs3$ -Promoter (Römer et al., 2007) erfolgt die transkriptionelle Aktivierung von  $Bs3_P\_EBE_{AvrBs4}Bs4C-R$  nicht nur durch AvrBs3 sondern auch durch AvrBs4, wie anhand der GUS-Färbungen in Abb. 13 zu erkennen ist. Diese Analysen belegen, dass allein das  $EBE_{AvrBs4}Bs4C-R$  notwendig und ausreichend für die Vermittlung der AvrBs4-Erkennung ist.

Im Vergleich der Promotoren von  $Bs4C-R$  und  $Bs4C-S$  wird deutlich, dass die zum  $EBE_{AvrBs4}Bs4C-R$  korrespondierende Sequenz in  $Bs4C-S$ , als  $EBE_{AvrBs4}Bs4C-S$  bezeichnet, an Position zwei und drei Nukleotidpolymorphismen aufweist (Abb. 14, siehe A6). Das  $EBE_{AvrBs4}Bs4C-S$  wurde ebenfalls in den  $Bs3$ -Promoter inseriert



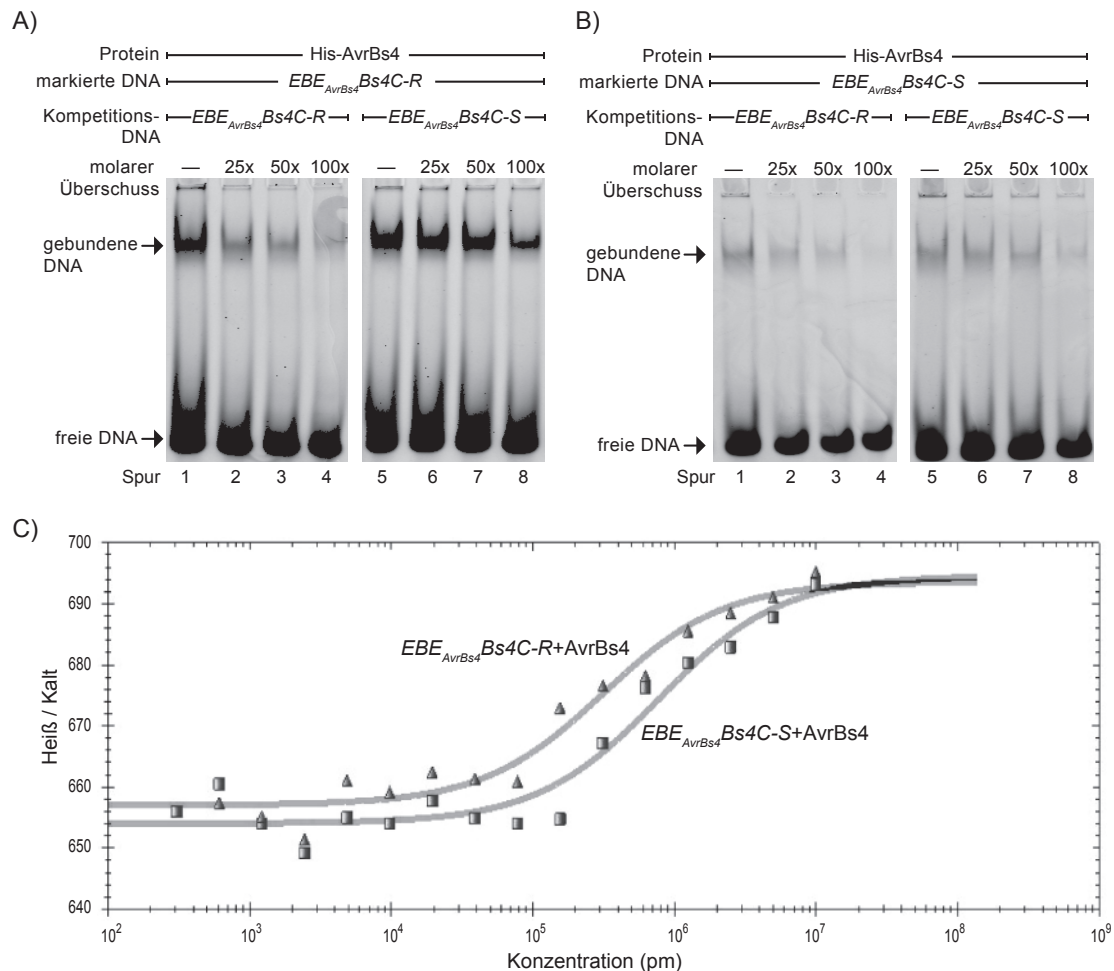
(*Bs3P\_EBE<sub>AvrBs4</sub>Bs4C-S*) und mittels GUS-Analyse auf seine Funktionalität hinterfragt. Die GUS-Analyse macht deutlich, dass *Bs3P\_EBE<sub>AvrBs4</sub>Bs4C-S* zwar durch AvrBs3, jedoch nicht durch AvrBs4 induziert wird (Abb. 13). Das lässt darauf schließen, dass die zwei zusätzlichen Nukleotidpolymorphismen im *EBE<sub>AvrBs4</sub>Bs4C-S* die AvrBs4-Erkennung verhindern.

### 3.7 AvrBs4 bindet mit höherer Affinität an das *EBE<sub>AvrBs4</sub>Bs4C-R* im Vergleich zum *EBE<sub>AvrBs4</sub>Bs4C-S*

Die beiden Polymorphismen zwischen dem funktionalen *EBE<sub>AvrBs4</sub>Bs4C-R* und dem nicht-funktionalen *EBE<sub>AvrBs4</sub>Bs4C-S* sind an Position zwei und drei zu finden (Abb. 14, siehe A6). Während das *EBE<sub>AvrBs4</sub>Bs4C-R* die Nukleotide T und A aufweist, enthält das *EBE<sub>AvrBs4</sub>Bs4C-S* zwei C-Nukleotide. Vergleicht man nun, die anhand des TALE-Codes vorhergesagten Nukleotide für die korrespondierenden AvrBs4-RVDs mit den entsprechenden Nukleotiden in den jeweiligen Effektorbindeelementen, wird deutlich, dass das *EBE<sub>AvrBs4</sub>Bs4C-R* die bevorzugten Nukleotide an dieser Position aufweist, verglichen zum *EBE<sub>AvrBs4</sub>Bs4C-S* (Abb. 14). Dies könnte in einer höheren Affinität von AvrBs4 gegenüber dem *EBE<sub>AvrBs4</sub>Bs4C-R* verglichen zum *EBE<sub>AvrBs4</sub>Bs4C-S* resultieren. Um dies zu hinterfragen, wurden in Kooperation mit der Arbeitsgruppe von Martin Parniske von Sylvia Singh *Electrophoretic Mobility Shift Assay* (EMSA)-Studien mit beiden Effektorbindeelementen durchgeführt (Abb. 15A+B) (Strauß et al., 2012, Fig. S10). Dafür wurde AvrBs4 aus *E. coli* aufgereinigt und mit fluoreszenzmarkierten *Bs4C-R*- bzw. *Bs4C-S*-Promoterfragmenten (Oligonukleotide) inkubiert und über eine native *Polyacrylamide Gel Electrophoresis* (PAGE) aufgetrennt. Die Bindung von AvrBs4 an das Oligonukleotid führt zu einer Änderung des Laufverhaltens (Verlangsamung) des Protein-DNA-Komplexes im Gel im Vergleich zum freien Protein – dem sogenannten „shift“. Um die Spezifität der Protein-DNA-Bindung zu bestätigen, erfolgt die Konkurrenz mit nicht-markierten Oligonukleotiden. Eine spezifische Protein-DNA-Bindung resultiert in einer Abnahme des Signals bei zunehmender Menge an nicht-markierten Oligonukleotiden, da die markierten Oligonukleotide durch die nicht-markierten verdrängt werden. Ist die Bindung nicht spezifisch, findet keine Verdrängung statt und somit kann keine Abnahme des Signals beobachtet werden. Wie in Abb. 15A Spur 1 an dem „shift“ zwischen freier und gebundener DNA zu erkennen ist, bindet AvrBs4 *in vitro* an das *EBE<sub>AvrBs4</sub>Bs4C-R*. Die Konkurrenz mit zunehmender Menge an nicht-markierten

*EBE<sub>AvrBs4</sub>Bs4C-R*-Oligonukleotiden führt zur Abnahme der Bandenintensität in Spur 2 bis 4 in Abb. 15A. Sie resultiert aus der Verdrängung der gebundenen markierten *EBE<sub>AvrBs4</sub>Bs4C-R*-DNA durch die nicht-markierte *EBE<sub>AvrBs4</sub>Bs4C-R*-DNA. Wird nicht-markierte *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA als Kompetitor verwendet (Abb. 15A Spur 6-8), nimmt die Bandenintensität kaum ab. Die Konkurrenzexperimente bestätigen somit die spezifische Bindung zwischen AvrBs4 und dem *EBE<sub>AvrBs4</sub>Bs4C-R*. AvrBs4 kann ebenfalls, wenn auch nur schwach, an markierte *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA binden (Abb. 15B Spur 1+5). Die Verdrängung der gebundenen *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA durch Zugabe nicht-markierter *EBE<sub>AvrBs4</sub>Bs4C-R*-DNA (Abb. 15B Spur 2-4) führt zu einer starken Abnahme der Bandenintensität bis hin zum Verlust. Eine Konkurrenz mit nicht-markierter *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA (Abb. 15B Spur 6-8) hingegen, führt nur zu einer geringen Änderung der Bandenintensität. Dies deutet auf eine geringere Affinität zwischen AvrBs4 und dem *EBE<sub>AvrBs4</sub>Bs4C-S* hin. Die Ergebnisse der EMSA-Studien zeigen deutlich, dass AvrBs4 eine höhere Affinität für das *EBE<sub>AvrBs4</sub>Bs4C-R* im Vergleich zum *EBE<sub>AvrBs4</sub>Bs4C-S* aufweist. Unterstützt wird dieses Resultat von den Ergebnissen der *microscale thermophoresis* (MST), eine Methode, die zur Affinitätsbestimmung zwischen AvrBs4 und dem *EBE<sub>AvrBs4</sub>Bs4C-R* bzw. dem *EBE<sub>AvrBs4</sub>Bs4C-S* genutzt wurde. Die MST-Analysen wurden von Christina Wolf aus der Arbeitsgruppe von Prof. Thomas Lahaye durchgeführt (Abb. 15C) (Strauß et al., 2012, Fig. S10). Hierfür wurde AvrBs4-Protein mit den für die EMSA-Analysen verwendeten fluoreszenzmarkierten *EBE<sub>AvrBs4</sub>Bs4C-R*- bzw. *EBE<sub>AvrBs4</sub>Bs4C-S*-Promoterfragmenten (Oligonukleotide) inkubiert und anschließend die Affinitäten mittels Thermophorese gemessen. Für die Bestimmung der Dissoziationskonstanten wurde die Menge an nicht-markierter *EBE<sub>AvrBs4</sub>Bs4C-R*- bzw. *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA variiert. Die errechneten Dissoziationskonstanten lagen bei 18,1 nM für die Interaktion AvrBs4 und *EBE<sub>AvrBs4</sub>Bs4C-R*-DNA sowie 181,5 nM für AvrBs4 und *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA und zeigen deutlich, dass das *EBE<sub>AvrBs4</sub>Bs4C-R* eine zehnfach höhere Affinität gegenüber AvrBs4 aufweist als das *EBE<sub>AvrBs4</sub>Bs4C-S* (Abb. 15C) (Strauß et al., 2012, Fig. S10).





**Abb. 15 EMSA und EMSA-Kompetitionsstudien sowie MST-Analysen zeigen, dass AvrBs4 eine höhere Affinität zum *EBE<sub>AvrBs4</sub>Bs4C-R* im Vergleich zum *EBE<sub>AvrBs4</sub>Bs4C-S* aufweist**

A+B) Es wurden EMSA und EMSA-Kompetitionsstudien für die Analyse der DNA-Bindung zwischen AvrBs4 und dem *EBE<sub>AvrBs4</sub>Bs4C-R* (A) bzw. dem *EBE<sub>AvrBs4</sub>Bs4C-S* (B) von Sylvia Singh durchgeführt. Dafür erfolgte die Inkubation von 5' DY682- (*EBE<sub>AvrBs4</sub>Bs4C-R*) oder 5' DY782- (*EBE<sub>AvrBs4</sub>Bs4C-S*) markierter DNA, die die jeweiligen EBE überspannen, mit 5 pmol His-AvrBs4-Fusionsprotein. Für die Konkurrenz und den Nachweis der spezifischen Bindung wurde ein molarer Überschuss von 25x, 50x und 100x nicht-markierter *EBE<sub>AvrBs4</sub>Bs4C-R*- und *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA eingesetzt. Die Position der freien DNA und der gebundenen DNA ist mit einem Pfeil markiert.

C) Für die MST-Analysen, welche von Christina Wolf durchgeführt wurden, wurde nicht-markierte DNA gegen eine feste Konzentration markierter DNA (25 nM) und AvrBs4 (400 nM) titriert. Dabei repräsentieren Dreiecke *EBE<sub>AvrBs4</sub>Bs4C-R*-DNA und Rechtecke *EBE<sub>AvrBs4</sub>Bs4C-S*-DNA. Aus den erhaltenen Kurven konnte für die Interaktion von AvrBs4 und *EBE<sub>AvrBs4</sub>Bs4C-R* eine Dissoziationskonstante von 18,1 nM und für AvrBs4 und *EBE<sub>AvrBs4</sub>Bs4C-S* eine Dissoziationskonstante 181,5 nM ermittelt werden.

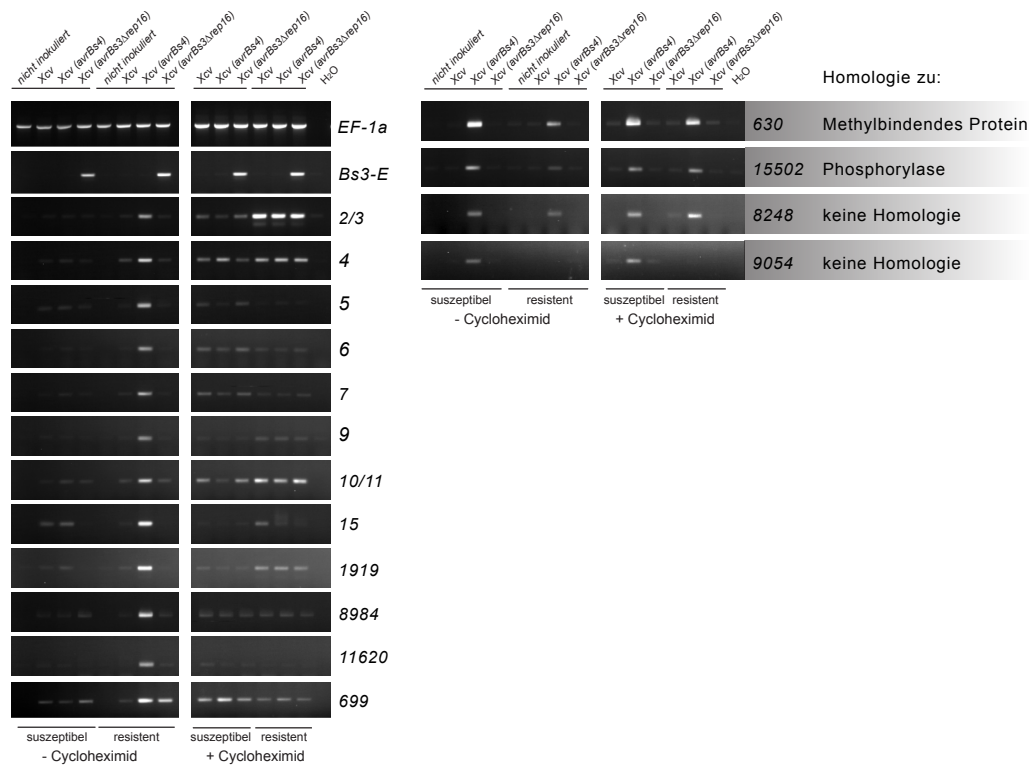
Abbildung modifiziert nach Strauß et al., 2012 Fig. S10. Die EMSA-Analysen wurden von Sylvia Singh (Abb. 15A+B) aus der Arbeitsgruppe von Prof. Martin Parniske und die MST-Analysen (Abb. 15C) von Christina Wolf aus der Arbeitsgruppe von Prof. Thomas Lahaye durchgeführt.

### 3.8 Analyse weiterer AvrBs4-induzierter Kandidaten

Wie bereits erwähnt (siehe 3.1, Strauß et al., 2012), wurden die Transkriptomanalysen zur Identifizierung von *Bs4C-R* in Kooperation mit der Firma Keygene von Remco van Poecke und Marcel Prins durchgeführt. Dabei wurden insgesamt 68 differentiell

AvrBs4-induzierte Gene identifiziert (siehe A1+A2, Strauß et al., 2012). Für 26 dieser Kandidaten lagen zusätzliche Sequenzinformationen vor. Diese stammen aus 454-Transkriptanalysen, die zu einem früheren Zeitpunkt initiiert wurden und bei denen AvrBs4-induziertes *C. pubescens*-Material 12 Stunden nach Infektion geerntet wurde (A2). Der Vorteil der 454-Sequenzierung liegt in den längeren Transkriptsequenzen (400-500 nt), die es ermöglichen Oligonukleotide zur Analyse der AvrBs4-induzierten Kandidaten mittels sq RT-PCR zu erstellen. Die Analyse der AvrBs4-induzierten Kandidaten erfolgte vor dem Hintergrund, dass die durchgeführten Transkriptanalysen neben dem Resistenzgen *Bs4C-R* auch potentielle Virulenztargets von AvrBs4 identifizieren sollten. Als Virulenztarget von AvrBs4, werden Gene bezeichnet, deren Expression die Vermehrung und Ausbreitung von *avrBs4*-exprimierenden Xanthomonaden begünstigen bzw. für die Ausbildung von Krankheitssymptomen entscheidend sind. Dies heißt, die AvrBs4-induzierten Kandidaten aus den Transkriptomanalysen können potentielle Virulenzzielgene darstellen. Allerdings sind bislang nur wenige TAL-induzierte Suszeptibilitätsgene bekannt (siehe 1.5). Der Großteil der AvrBs4-induzierten Gene werden höchstwahrscheinlich sogenannte *off-targets* repräsentieren, die zwar EBE in ihren Promotoren aufweisen aber nicht von Relevanz für die bakterielle Virulenz sind.

Mit den 25 Kandidaten, für welche zusätzliche Sequenzinformationen vorhanden waren, wurden sq RT-PCR-Analysen in An- und Abwesenheit von CHX durchgeführt. Dies ermöglicht die Differenzierung zwischen direkten und indirekten AvrBs4-induzierten Genen. Die sq RT-PCR-Analysen zeigten, dass fünf Kandidaten durch *Xanthomonas* allein induziert waren. Desweiteren konnten 16 Kandidaten identifiziert werden, deren Expression in der resistenten *C. pubescens*-Akzession induziert war. Die Analysen mit CHX-behandeltem Material lassen jedoch vermuten, dass nicht alle direkte Zielgene von AvrBs4 sind (Abb. 16). Für vier Kandidaten konnte bestätigt werden, dass sie direkt durch AvrBs4 induziert werden, da eine Transkriptanreicherung auch in den CHX-behandelten Proben nachgewiesen werden konnte. Von diesen vier Kandidaten werden drei (Kandidaten 639, 8248 und 15502, Abb. 16) in beiden *C. pubescens*-Akzessionen und einer ausschließlich in der suszeptiblen Akzession durch AvrBs4 induziert (Kandidat 9054, Abb. 16). Die in beiden Akzessionen durch AvrBs4 direkt-induzierten Kandidaten stellen die vielversprechendsten potentiellen Virulenzgene dar und deshalb wurden diese drei Kandidaten intensiver untersucht.



**Abb. 16 Auswahl einiger sq RT-PCR-Analysen in An- und Abwesenheit von CHX zur Bestimmung von AvrBs4-direkt induzierten Genen**

*C. pubescens* Blätter wurden mit *Xcv*-Stämmen, die entweder keine oder TALE-Gene exprimieren (AvrBs4 und AvrBs3Arep16), mit und ohne (+/-) Cycloheximid, infiziert und 18 Stunden nach Inokulation geerntet. Sq RT-PCR-Analysen wurden mit Oligonukleotiden zur Amplifikation der verschiedenen Kandidatengene durchgeführt (siehe A34). Als Kontrolle für die Verwendung vergleichbarer cDNA-Mengen wurden Oligonukleotide für den Elongationsfaktor 1 $\alpha$  (*EF-1 $\alpha$* ) verwendet. Die erhaltenen PCR-Produkte wurden auf einem 1,5 %igen Agarosegel aufgetrennt und mit Ethidiumbromid angefärbt.

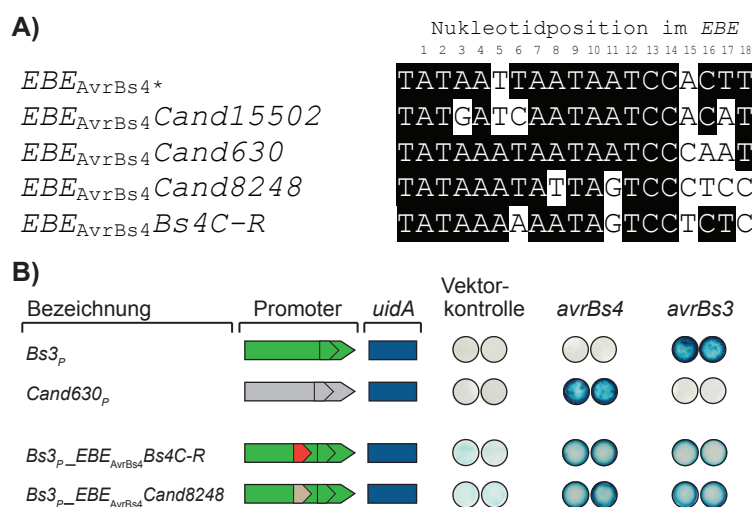
### 3.8.1 Kandidat 630

Für das Kandidatengen 630 standen durch das erste Transkriptomprofiling Sequenzinformationen von etwa 500 nt zur Verfügung. Da keine vollständige KDS vorlag, wurde eine *C. pubescens*-cDNA-Bibliothek (AvrBs4-induziert, 12 hpi, siehe 2.14) mit spezifischen Oligonukleotiden für den Kandidaten 630 durchsucht. Es konnte ein Einzelklon isoliert, sequenziert und die KDS für den Kandidaten 630 bestimmt werden (siehe A7). Jedoch gab es keine Informationen über den Promoterbereich. Durch die Zusammenarbeit mit der Arbeitsgruppe, die das *C. annuum* cv. CM334-Genom sequenziert (Prof. Doil Choi) (Kim et al., 2014), konnte ein größerer Genomabschnitt identifiziert werden, der neben der KDS auch Promotersequenzen für den Kandidaten 630 aufwies. Aufgrund dieser Sequenzinformationen konnten Oligonukleotide abgeleitet werden und die Sequenz in den *C. pubescens*-Akzessionen PI 235047 und PI 585270 bestimmt werden (siehe A7). Beide *C. pubescens* 630-Sequenzen, sowohl Promoter- als auch KDS, sind

identisch und zeigen drei AS-Polymorphismen verglichen zur *C. annuum*-Sequenz für den Kandidaten 630 (siehe A8).

Das *C. pubescens*-Gen des Kandidaten 630 besteht aus zwei Exons und einem Intron und kodiert für ein 173 AS großes Protein, welches Homologien zu Vertretern der Familie der Methyl-CpG-bindenden Proteinen (MBP) aufweist (siehe A7). Während ein Methyl-CpG-bindendes Protein aus Soja (*Glycine max*) auf Aminosäureebene zu 56 % identisch ist, zeigt eines aus Kartoffel (*Solanum tuberosum*) eine Identität von 84 % auf Nukleotidebene (siehe A9+A10).

Da der Kandidat 630 als direktes Zielgen von AvrBs4 bestätigt wurde (siehe 3.8, Abb. 16), konnte im Promoterbereich ein potentiell EBE identifiziert werden, welches 62 Nukleotide 3' vom Startkodon liegt. Der Vergleich mit dem TALE-Code vorhergesagten EBE ist in Abb. 17A dargestellt. Das  $EBE_{AvrBs4}Cand630$  weist vier Polymorphismen verglichen zu dem  $EBE_{AvrBs4}^*$  auf, wobei sich drei im 3'-Bereich an Position 15, 16 und 17 befinden. Der vierte konnte an Position fünf identifiziert werden und entspricht dem, der auch in dem  $EBE_{AvrBs4}Bs4C-R$  vorhanden ist (Abb. 17A). Um den Promoter des Kandidaten 630 funktional zu testen, wurde ein 250 nt umfassendes Promoterfragment vor ein *uidA*-Gen kloniert und transient in *N. benthamiana* mit AvrBs4 koexprimiert. Wie in Abb.17B anhand der blau gefärbten Blattscheiben zu erkennen ist, induziert AvrBs4 den Promoter des Kandidaten 630. Dies zeigt, dass der Kandidat 630 für ein Gen kodiert, welches ein potentiell Virulenzgen darstellen könnte, da es direkt und in beiden Akzessionen durch AvrBs4 induziert wird.



**Abb. 17 Vergleich und funktionaler Test der EBE AvrBs4-induzierter Gene**

A) Sequenzvergleich der EBE weiterer AvrBs4-induzierter Gene (*Cand15502*, *Cand630*, *Cand8248*) mit dem anhand des TALE-Code vorhergesagten  $EBE_{AvrBs4}^*$  und dem  $EBE_{AvrBs4}Bs4C-R$ . Weiße Buchstaben auf schwarzem Hintergrund markieren identische Nukleotide.

B) GUS-Analysen nach *Agrobacterium*-vermittelter Transformation von *N. benthamiana* mit Promoter-*uidA*-T-DNA-Konstrukten in Kombination mit leerem Vektor, den TALE-Genen *avrBs4* und *avrBs3*. Das Promoterfragment des Kandidaten 630 ist grau und das, als Kontrolle dienende *Bs3*-Promoterfragment, in Grün dargestellt. Pfeile innerhalb der Promotoren markieren die EBE. Blattscheiben wurden 38 Stunden nach Infiltration geerntet, mit X-Gluc gefärbt und in Ethanol gebleicht um die Blaufärbung hervorzuheben.

### 3.8.2 Kandidat 15502

Die KDS des Kandidaten 15502 aus *C. annuum* cv. CM334 beläuft sich auf etwa 9 kb. Mit Hilfe des Vorhersageprogramms (Softberry) konnte ein offener Leserahmen von etwa 3 kb bestimmt werden (siehe A11+A12), der aus 15 Exons und 14 Introns besteht. Kandidat 15502 kodiert für ein Protein, welches Homologien zu einer Phosphorylase aufweist (siehe A13). Dabei zeigt eine alpha-Glucan-Phosphorylase aus Kartoffel (*Solanum tuberosum*) eine Sequenzidentität von 90,56 % (siehe A14). Alpha-Glucan-Phosphorylasen spielen eine wichtige Rolle im Kohlenhydrat-Stoffwechsel. Außerdem sind Zucker nicht nur in metabolischen Prozessen in der Pflanze involviert sondern auch in Signalprozessen (Ruan, 2014).

Im Promoterbereich wurde eine potentiell EBE für AvrBs4 identifiziert, deren Vergleich mit dem vorhergesagten  $EBE_{AvrBs4*}$  in Abb. 17A zeigt, dass das  $EBE_{AvrBs4}Cand15502$  nur drei Polymorphismen zu diesem aufweist. Anstatt eines für das RVD NI vorhergesagte A- befindet sich an Position drei ein G-Nukleotid. Position sechs und 17 weisen ein C- und ein A-Nukleotid auf. Jedoch bevorzugen die dort präsenten RVDs NG ein T-Nukleotid (Abb. 17A). Das  $EBE_{AvrBs4}Cand15502$  wurde noch nicht in transienten GUS-Studien getestet. Die vergleichende Sequenzierung stellte sich aufgrund der Größe des Kandidaten 15502 als sehr schwierig heraus. Bislang konnte die KDS aus beiden Akzessionen nur teilweise bestimmt werden, diese wies keine Unterschiede zwischen suszeptibler und resistenter *C. pubescens*-Akzession auf. Die cDNA beider Akzessionen konnte derweil komplett sequenziert werden und zeigte ebenfalls keine Abweichungen voneinander (siehe A12).

### 3.8.3 Kandidat 8248

Für den Kandidaten 8248 konnte ebenfalls ein potentiell EBE für AvrBs4 identifiziert werden, dessen Sequenzvergleich mit dem TALE-Code vorhergesagten  $EBE_{AvrBs4*}$  und dem  $EBE_{AvrBs4}Bs4C-R$  in Abb. 17A dargestellt ist (siehe A15). Der Sequenzvergleich macht deutlich, dass das  $EBE_{AvrBs4}Cand8248$  die größte Zahl an Abweichungen (sieben) zu dem vorhergesagten  $EBE_{AvrBs4*}$  aufweist, jedoch dem

funktionalen *EBE<sub>AvrBs4Bs4C-R</sub>* am ähnlichsten ist. Der funktionale Test des *EBE<sub>AvrBs4Cand8248</sub>* erfolgte anhand von GUS-Analysen und zeigte, dass auch dieses EBE AvrBs4-Erkennung vermittelt (Abb. 17B).

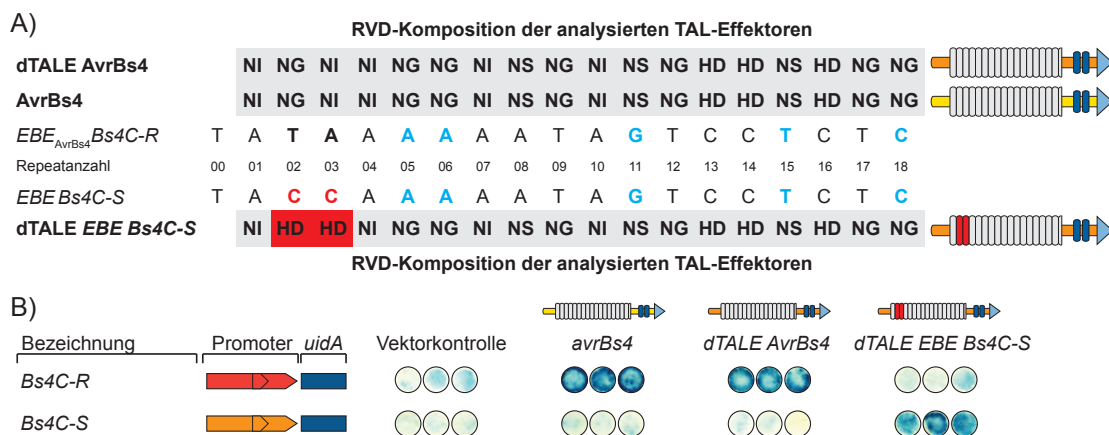
Zu Bearbeitungsbeginn lagen für den Kandidaten 8248 Sequenzinformationen von etwa 500 nt vor. Über das *C. annuum* cv. CM334-Genomprojekt (siehe A15) konnten zusätzliche Sequenzinformationen erhalten werden, die anschließend mit Hilfe von Genvorhersage-Programmen (Softberry) bearbeitet wurden. Aber es war nicht möglich, eine KDS zu identifizieren bzw. vorherzusagen. Die Suche in cDNA- und genomischen Datenbanken lieferte keine homologen Sequenzen weder auf Nukleotid- noch auf Proteinebene. Für das identifizierte EBE konnte aber gezeigt werden, dass es durch AvrBs4 induziert wird. Möglicherweise liegt die KDS des Kandidaten 8248 weiter *downstream* von seinem EBE. Jedoch sind die meisten TALE-induzierten-EBEs in einem Abstand von 40-130 nt *upstream* zum Transkriptionsstart zu finden (Antony et al., 2010; Hummel et al., 2012; Kay et al., 2007; Kay et al., 2009; Römer et al., 2007; Römer et al., 2009a; Römer et al., 2009b; Strauß et al., 2012). In diesem Bereich konnte jedoch kein Startkodon bzw. eine offener Leserahmen identifiziert werden. Wenn die Nukleotidsequenz des EBE des Kandidaten 8248 zufällig ein passendes EBE für AvrBs4 darstellt, kann eine Transkription initiiert werden, muss jedoch nicht in der Translation der transkribierten RNA resultieren. Somit kann Kandidat 8248 in den Transkriptanalysen zwar detektiert werden, ohne jedoch in ein Protein translatiert zu werden.

### 3.9 Assemblierung eines designer TALE zur Induktion von *Bs4C-S*

Bislang konnte gezeigt werden, dass *Bs4C-R* und *Bs4C-S* im Bereich des EBE zwei Nukleotidpolymorphismen aufweisen, die bedingen, dass nur *Bs4C-R*, nicht jedoch *Bs4C-S* durch AvrBs4 induziert wird (siehe 3.6, Abb. 13). Allerdings resultiert die Überexpression von *Bs4C-S* in einer HR, was zeigt, dass das Protein funktional ist (siehe 3.5, Abb. 12). Es stellte sich daher die Frage, ob eine AvrBs4-Mutante, welche in Repeat zwei und drei, die zur Sequenz im *Bs4C-S*-Promoter (nachfolgend als *EBE Bs4C-S* bezeichnet) korrespondierenden RVDs aufweist (Abb. 18) spezifisch den *Bs4C-S*- und nicht den *Bs4C-R*-Promoter induzieren kann.

In unserem Labor wurde ein „Baukasten“ (tool kit) etabliert, der es uns ermöglicht, schnell und einfach TAL-Effektorgene, kodierend für TAL-Proteine mit gewünschter

RVD-Komposition, sogenannte designer TALEs (dTALEs) zu erstellen (siehe 1.4) (Morbitzer et al., 2011). Zur Induktion des *EBE Bs4C-S* wurde der dTALE *EBE Bs4C-S* von Janett Elsaesser aus der Arbeitsgruppe von Prof. Thomas Lahaye generiert (Abb. 18, Janett Elsaesser, unveröffentlichte Daten). Dieser wurde anschließend in Promoter-*uidA*(GUS)-Studien in *N. benthamiana* getestet (Abb. 18). Dafür wurde der dTALE transient via *Agrobacterium* zusammen mit *Bs4C-R*- bzw. *Bs4C-S*-Promoter-*uidA*-Fusionen *in planta* koexprimiert. Wie in der Abb. 18 zu erkennen, induzierte der dTALE *EBE Bs4C-S* den *Bs4C-S*-Promoter, aber nicht den *Bs4C-R*-Promoter. Als Kontrolle wurden parallel *avrBs4* und ein dTALE mit AvrBs4 RVD-Komposition (dTALE AvrBs4) zusammen mit *Bs4C-R*- und *Bs4C-S*-Promoter-*uidA*-Fusionen kotransformiert. Beide TAL-Effektoren induzierten den *Bs4C-R*-, nicht aber den *Bs4C-S*-Promoter. Zusammengefasst konnte gezeigt werden, dass zwei Nukleotide bzw. zwei RVDs entscheidend für die spezifische, transkriptionelle Induktion der *Bs4C-R*- bzw. *Bs4C-S*-Promotoren sind.



**Abb. 18 Der dTALE *EBE Bs4C-S* aktiviert spezifisch den *Bs4C-S*-Promoter aber nicht den *Bs4C-R*-Promoter**

A) Dargestellt ist die RVD-Abfolge von *avrBs4* sowie der assemblierten dTALEs, dTALE *avrBs4* und dTALE *EBE Bs4C-S* und die korrespondierenden Nukleotide des *EBE<sub>AvrBs4</sub>Bs4C-R* und des *EBE Bs4C-S*. Identische RVDs sind grau hinterlegt und RVD-Unterschiede des dTALE *EBE Bs4C-S* sind rot markiert. Blaue Buchstaben kennzeichnen Nukleotide, die nicht dem TALE-Code entsprechen.

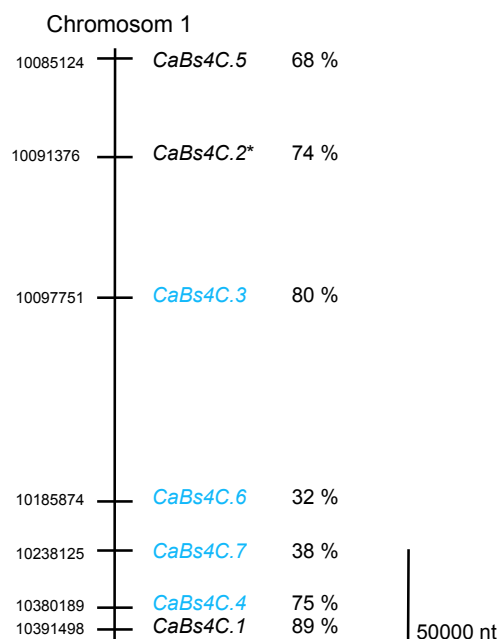
B) GUS-Analysen nach *Agrobacterium*-vermittelter Transformation von *N. benthamiana* mit Promoter-*uidA*-T-DNA-Konstrukten in Kombination mit leerem Vektor, *avrBs4* und den erstellten dTALEs (dTALE *avrBs4* und dTALE *EBE Bs4C-S*). *AvrBs4* und der dTALE *avrBs4* besitzen zwar die gleiche RVD-Abfolge, unterscheiden sich aber in der Komposition des N- und C-Terminus (gelb und orange dargestellt) aufgrund der Konstruktion der dTALEs (siehe 2.12.14). Blattscheiben wurden 40 Stunden nach Infiltration geerntet, mit X-Gluc gefärbt und abschließend in Ethanol gebleicht.

### 3.10 Analyse von *Bs4C*-Homologen in *Capsicum annuum*

#### 3.10.1 Identifizierung von *Bs4C*-Homologen im assemblierten Genom von *C. annuum* cv. CM334

Während der Arbeiten an *Bs4C-R* (*C. pubescens*) erfolgte in der Arbeitsgruppe von Professor Doil Choi (Seoul National University, Südkorea) die Sequenzierung und Assemblierung des Genoms von *C. annuum* cv. CM334 (Kim et al., 2014). In Zusammenarbeit mit der Arbeitsgruppe von Professor Doil Choi konnten, in der bis dahin assemblierten Genomsequenz von CM334, sieben *Bs4C*-Orthologe identifiziert werden, die auf Nukleotidebene 32-89 % identisch zu *Bs4C-R* (korrespondierende Sequenzen von Start- bis Stoppkodon) sind (Tab. 2, siehe A16). Die Nummerierung der *Bs4C*-Orthologen erfolgte dabei nach dem Zeitpunkt ihrer Identifikation.

Die Orthologen konnten auf Chromosom 1 lokalisiert werden und sind nachfolgend als *CaBs4C.1* bis *CaBs4C.7* bezeichnet (Tab. 2). Die Abb. 19 zeigt die Position und die relative Anordnung der *Bs4C*-Orthologen zueinander auf Chromosom 1 von *C. annuum* cv. CM334.



**Abb. 19 Anordnung der *Bs4C-R*-Homologen auf Chromosom 1 des *C. annuum* cv. CM334-Genoms**

Auf dem Chromosom 1 des *C. annuum* cv. CM334-Genoms wurden die sieben *Bs4C-R*-Homologen und ihre Positionen dargestellt. Dabei sind die Sequenzidentitäten verglichen zur *Bs4C-R*-Nukleotidsequenz in Prozent angegeben. Hellblau unterlegt sind Homologe, welche aufgrund von Insertions- und Deletionspolymorphismen keinen mit *Bs4C-R* vergleichbaren durchgehenden Leserahmen aufweisen.



### 3.11 Analyse der *Bs4C*-Homologen *CaBs4C.1* bis *CaBs4C.7*

#### 3.11.1 ORF und ORF-korrespondierende Sequenzen

Die Tab. 2 fasst die wichtigsten Informationen zu den Orthologen *CaBs4C.1-7* aus *C. annuum* cv. CM334 zusammen. Der Sequenzvergleich der sieben Orthologen mit *Bs4C-R* machte deutlich, dass *CaBs4C.1* die höchste Homologie (89 %) zu *Bs4C-R* aufweist. Außerdem besitzt es einen durchgängigen Leserahmen (Tab. 2, siehe A16). Der *CaBs4C.1*- und *Bs4C-R*-Sequenzvergleich zeigte insgesamt 21 AS- und 8 InDel-Polymorphismen auf, die bedingen, dass beide Proteine über eine Homologie von 81 % verfügen (Abb. 21A, Tab. 2). Neben *CaBs4C.1* besitzen auch *CaBs4C.2* und *CaBs4C.5* einen durchgängigen Leserahmen, während *CaBs4C.3*, *CaBs4C.4*, *CaBs4C.6* und *CaBs4C.7* aufgrund von Insertions- und Deletionspolymorphismen keinen mit *Bs4C-R* vergleichbaren Leserahmen zeigen (Tab. 2, siehe A16). Das *C. annuum* Ortholog *CaBs4C.2* besitzt 23 AS mehr verglichen zu *CaBs4C.1*. Eine Sequenzanalyse konnte eine Duplikation eines 26 AS-Motivs an Position 52 in *CaBs4C.2* identifizieren (Tab. 2, siehe A17). Zusätzlich konnten zwischen *CaBs4C.2* und *Bs4C-R* insgesamt 35 Polymorphismen (AS- und InDel-Polymorphismen) ermittelt werden, welche in einer Homologie von 69,3 % resultieren (Tab. 2, siehe A16+A17). Das Homolog *CaBs4C.5* hingegen ist 10 AS kürzer als *Bs4C-R* und weist mit insgesamt 75 die höchste Zahl an Polymorphismen (AS- und InDel-Polymorphismen) auf (siehe A16+A18). Die Homologie zu *Bs4C-R* beträgt 54,2 % (Tab. 2, siehe A16).

Indes zeigt *CaBs4C.3* in den ersten 50 Nukleotiden einen Deletionspolymorphismus von 14 nt und einen Insertionspolymorphismus von 4 nt auf, was zu einem vorzeitigen Stoppkodon und einer KDS von nur 24 nt führt (Tab. 2, siehe A16). Bei *CaBs4C.4* resultiert die Insertion eines T-Nukleotids an Position acht in der Verschiebung des Leserahmens, was ein vorzeitiges Stoppkodon nach 12 nt zur Folge hat (Tab. 2, siehe A16). Erstaunlicherweise weisen sowohl *CaBs4C.3* als auch *CaBs4C.4* ab ihrem zweiten ATG in der Nukleotidsequenz durchgehende Leserahmen und kodieren somit für verkürzte *Bs4C-R*-homologe Sequenzen von 127 bzw. 147 AS auf (siehe A19).

**Tab. 2 Identifikation von *Bs4C-R*-Homologen in *C. annuum* cv. CM334 und deren Vergleich mit *Bs4C-R* und *Bs4C-S***

Name	zu <i>Bs4C-R</i> korresp. Sequenz <sup>1)</sup>	KDS <sup>2)</sup> in nt	Homologie <sup>3)</sup> zu				Sequenz- vergleich zu <i>C. a.</i> ECW/30R <sup>6)</sup>	35S-Über- expression
			<i>Bs4C-R</i> <sup>4)</sup>	<i>Bs4C-R</i> <sup>5)</sup>	<i>Bs4C-S</i> <sup>4)</sup>	<i>Bs4C-S</i> <sup>5)</sup>		
<i>Bs4C-R</i>	495 nt	495	100 %	100 %	95,6 %	93,3 %		HR
<i>Bs4C-S</i>	495 nt	495	95,6 %	93,3 %	100 %	100 %		HR
<i>CaBs4C.1</i>	495 nt	495	88,8 %	80,9 %	86,9 %	79,3 %	identisch	HR
<i>CaBs4C.2</i>	564 nt	564	73,7 %	69,3 %	74,7 %	70,9 %	identisch	keine HR <sup>7)</sup>
<i>CaBs4C.3</i>	491 nt	24	80,3 %	-	80,9 %	-	n.g.	n.g.
<i>CaBs4C.4</i>	503 nt	9	75 %	-	73,7 %	-	n.g.	n.g.
<i>CaBs4C.5</i>	465 nt	465	67,6 %	54,2 %	68,2 %	54,8 %	n.g.	n.g.
<i>CaBs4C.6</i>	351 nt	39	32,2 %	-	31,8 %	-	n.g.	n.g.
<i>CaBs4C.7</i>	492 nt	120	38,5 %	-	38,9 %	-	n.g.	n.g.

Grau hinterlegt sind die Homologen, die auf Chromosom 1 des *C. annuum* cv. CM334-Genom liegen (Seon-In Yeom, Seoul).

Sequenzen und Vergleiche sind im Anhang A16

<sup>1)</sup>Bezieht sich auf die *Bs4C-R*-Sequenz von Start- bis Stoppkodon.

<sup>2)</sup>Die potentielle KDS wurde mittels Vorhersageprogramm Softberry oder anhand von Sequenzvergleichen bestimmt.

<sup>3)</sup>Sequenzhomologie von *Bs4C-R*-KDS/*Bs4C-R* im Vergleich zu den *Bs4C-R*/*Bs4C-R*-korrespondierenden Sequenzen.

<sup>4)</sup>Der Vergleich bezieht sich auf die zur *Bs4C-R*/*Bs4C-S*-Sequenz korrespondierenden Sequenzen von potentiellen Start- bis Stoppkodon.

<sup>5)</sup>Es wurden nur Vergleiche mit Homologen durchgeführt, die einen durchgängigen Leserahmen aufweisen.

<sup>6)</sup>Der Vergleich bezieht sich auf die zur *Bs4C-R*-Sequenz korrespondierenden KDS.

<sup>7)</sup>Stephanie Solle, Forschungsgruppenpraktikum

KDS-kodierende Sequenz, HR-hypersensitive Reaktion, n.g.-nicht getestet, *C.a.-Capsicum annuum*

### 3.11.2 EBE in den Promotersequenzen von *CaBs4C.1*- *CaBs4C.5*

Neben den KDS der *C. annuum* cv. CM334-Homologen wurden auch die Sequenzen vor dem potentiellen Startkodon (etwa 300-400 Nukleotide) miteinander verglichen und es zeigte sich, dass diese recht divers sind (siehe A22). In den Promoterbereichen konnten durch Sequenzvergleiche mit dem anhand des TALE-Code vorhergesagten EBE (*EBE<sub>AvrBs4\*</sub>*) und Sequenzalignments (ClustalW) potentielle EBE für AvrBs4 in den Orthologen identifiziert werden (Tab. 3). Diese liegen in einem Bereich von 151 nt bis 243 nt *upstream* des potentiellen Startkodons (Tab.3).

**Tab. 3 Identifikation<sup>1)</sup> und Vergleich potentieller EBE der *Bs4C-R*-Homologen aus *C. annuum* cv. CM334**

Name des potentiellen <i>EBE</i>	Abstand zum potentiellen ATG	Anzahl der PM zum <i>EBE<sub>AvrBs4*</sub></i>	Sequenz der <i>EBE</i> und Vergleich	funktionaler Test <sup>2)</sup> <i>AvrBs4</i> - Induzierbarkeit
<i>EBE<sub>AvrBs4*</sub></i>			TATTAATTAAATCCACTT-	ja
<i>EBE<sub>AvrBs4</sub>Bs4C-R</i>	162 nt	5	TATTAATAAATAGTCCTCTC-	ja
<i>EBE<sub>AvrBs4</sub>Bs4C-S</i>	155 nt	7	TACCAAAAATAGTCCTCTC-	nein
<i>EBE<sub>AvrBs4</sub>CaBs4C.1</i>	232 nt	9	TACCAAAAATAGCCCTCCC-	nein
<i>EBE<sub>AvrBs4</sub>CaBs4C.2</i>	214 nt	10	TGCCAAAAATAGCCCTCCA-	n.g.
<i>EBE<sub>AvrBs4</sub>CaBs4C.3</i>	176 nt	12	TGCCAAAAATAGCCCTACAT	n.g.
<i>EBE<sub>AvrBs4</sub>CaBs4C.4</i>	151 nt	11	TGCCAAAAATAGCCCTCCAT	n.g.
<i>EBE<sub>AvrBs4</sub>CaBs4C.5</i>	243 nt	12	TACCCGATATGTACACCTA-	n.g.

<sup>1)</sup>Die Identifikation der potentiellen EBE erfolgte durch Sequenzvergleiche, ClustalW- und T-Coffee-Analysen.

<sup>2)</sup> Insertion des potentiellen EBE in ein *Bs3*-Promoter-GUS-Konstrukt oder Klonierung eines Promoterfragments vor ein *uidA*-Gen und transiente Ko-Expression in *N. benthamiana* mit AvrBs4.

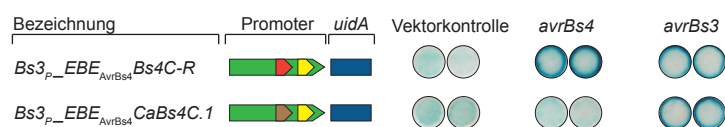
n.g. – nicht getestet, PM - Polymorphismen, *EBE<sub>AvrBs4\*</sub>* - anhand des TALE-Codes für AvrBs4-RVDs vorhergesagtes Effektorbindelement

Ein Vergleich der potentiellen EBE der *C. annuum* cv. CM334-Orthologen (Tab. 3) legte offen, dass keines ein EBE aufweist, das dem von *Bs4C-R* entspricht und deshalb wahrscheinlich nicht durch AvrBs4 induziert werden kann. Außerdem zeigten alle putativen EBE-Sequenzen aus *C. annuum* cv. CM334 an Position zwei und drei C-Nukleotide, wie das potentielle EBE von *Bs4C-S*, welches nachweislich nicht AvrBs4-induzierbar ist (siehe 3.6). Außerdem weisen die potentiellen EBE noch zwischen zwei und zehn zusätzlichen Polymorphismen verglichen zu dem  $EBE_{AvrBs4}Bs4C-R$  auf, was vermuten lässt, dass keines dieser vorhergesagten EBE durch AvrBs4 induziert werden kann. Dies stimmt mit den phänotypischen Analysen überein, die zeigten, dass AvrBs4 keine HR auf *C. annuum* cv. CM334 bzw. auf *C. annuum* ECW-30R auslösen kann (unveröffentlichte Daten T. Strauß, Schornack et al., 2005).

### 3.11.3 Funktionale Analyse der *Bs4C*-Homologen aus *C. annuum*

Für funktionale Analysen der *Bs4C*-Orthologen *CaBs4C.1* und *CaBs4C.2* wurden Promoter- und kodierende Fragmente aus den CM334 verwandten und uns zur Verfügung stehenden *C. annuum*-Kultivar ECW und ECW-30R amplifiziert und sequenziert. Vergleiche der KDS zeigten, dass *CaBs4C.1* aus *C. annuum* ECW bzw. ECW-30R und cv. CM334 zueinander identisch sind, während im Promoter eine Insertion zwischen EBE und Startkodon in *C. annuum* ECW bzw. ECW-30R identifiziert werden konnte (siehe A23).

Im Folgenden wurde das putative EBE von *CaBs4C.1* in den *Bs3*-Promoter kloniert und in GUS-Studien in *N. benthamiana* auf AvrBs4-Induzierbarkeit geprüft (Römer et al., 2009a). Das  $EBE_{AvrBs4}Bs4C-R$ , welches ebenfalls in den *Bs3*-Promoter integriert wurde, diente als Kontrolle (siehe 3.6). Wie in Abb. 20 anhand der blau gefärbten Blattscheiben zu erkennen ist, kann AvrBs4 nur das *Bs3*-Promoter-GUS-Konstrukt induzieren, welches das  $EBE_{AvrBs4}Bs4C-R$  ( $Bs3_P\_EBE_{AvrBs4}Bs4C-R$ ) jedoch nicht das  $EBE_{AvrBs4}CaBs4C.1$  ( $Bs3_P\_EBE_{AvrBs4}CaBs4C.1$ ) aufweist. Als Positivkontrolle diente AvrBs3, welches das in beiden Konstrukten vorhandene  $EBE_{AvrBs3}Bs3$  induzierte.

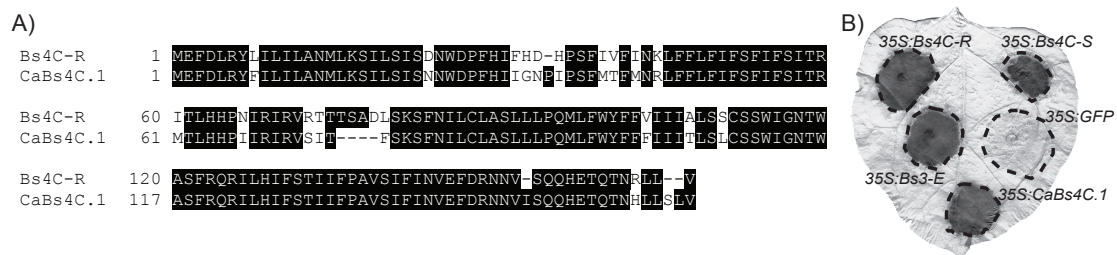


**Abb. 20 Das  $EBE_{AvrBs4}CaBs4C.1$  wird nicht durch AvrBs4 induziert**

GUS-Analysen nach *Agrobacterium*-vermittelter Transformation von *N. benthamiana* mit Promoter-*uidA*-T-DNA-Konstrukte in Kombination mit leerem Vektor, *35S:avrBs4* und *35S:avrBs3*.

Blattscheiben wurden 36 Stunden nach Infiltration geerntet, mit X-Gluc gefärbt und anschließend in Ethanol gebleicht.

Das putative EBE von *CaBs4C.1* ist, wie das von *Bs4C-S*, nicht durch AvrBs4 induzierbar (siehe 3.6). Allerdings zeigten Überexpressionsanalysen, dass *Bs4C-S*, trotz zahlreicher Polymorphismen im Vergleich zu *Bs4C-R*, funktional ist und in der Lage, eine HR zu induzieren (siehe 3.5). Um dies für *CaBs4C.1* zu hinterfragen, welches 29 Polymorphismen (Abb. 21A) verglichen zu *Bs4C-R* aufweist, wurde die KDS in einen Expressionsvektor unter Kontrolle eines 35S-Promoters kloniert und transient in *N. benthamiana* exprimiert (Abb. 21B). Da die Überexpression von *CaBs4C.1* in der Ausbildung einer HR resultierte, wie anhand der Abb. 21B zu erkennen ist, haben die Polymorphismen in der KDS von *CaBs4C.1* keine detektierbare funktionale Relevanz (Abb. 21A). Die Analysen von *CaBs4C.1* ergaben, dass es bei Überexpression zwar eine HR auslösen kann, jedoch kein funktionales AvrBs4-induzierbares EBE im Promoter aufweist.



**Abb. 21 Die konstitutive Expression des *C. annuum* *Bs4C*-Ortholog *CaBs4C.1* resultiert in einer HR.**

A) Der Vergleich der Proteinsequenzen von *Bs4C-R* und *CaBs4C.1* zeigt 21 AS-Polymorphismen sowie 8 InDel-Polymorphismen, welche weiß hinterlegt sind. Identische AS sind weiß auf schwarzem Hintergrund dargestellt.

B) Phänotypen nach *Agrobacterium*-vermittelter Transformation von *N. benthamiana*. Die KDS des *C. annuum*-Orthologs *CaBs4C.1* wurden unter Kontrolle des *CaMV*-35S-Promoters in einen T-DNA-Vektor kloniert und *Agrobacterium*-vermittelt in *N. benthamiana* exprimiert. Als Negativkontrolle diente ein 35S:*GFP*-, als Positivkontrolle ein 35S:*Bs3-E*-Konstrukt (Römer et al., 2007). Die Blätter wurden drei Tage nach Inokulation geerntet und zum besseren Sichtbarmachen der HR (dunkle Blattbereiche) in Ethanol entfärbt.

Sequenzanalysen zwischen den Kultivaren CM334 und ECW bzw. ECW-30R ergaben für *CaBs4C.2*, dass sie sowohl im Promoter als auch in der KDS identisch sind (siehe A24). Neben *CaBs4C.1* erfolgte auch die Analyse des *C. annuum*-Homologs *CaBs4C.2* in Überexpressionsstudien. Dabei konnte gezeigt werden, dass *CaBs4C.2* nicht in der Lage ist, eine HR bei Überexpression zu induzieren (Stephanie Solle, Forschungsgruppenpraktikum). Durch die Sequenzanalysen der verschiedenen *C. annuum*-Kultivare konnte gezeigt werden, dass *CaBs4C.2* eine Duplikation von 26

AS im Vergleich zu Bs4C-R aufweist (siehe 3.11.1, siehe A16). Der Einschub dieser Aminosäuresequenz kann einen Einfluss auf die Ausbildung der Proteinstruktur von CaBs4C.2 haben. Dass heißt, CaBs4C.2 ist möglicherweise nicht mehr in der Lage seine funktionale Struktur auszubilden. Dies müsste in weiteren Analysen durch Deletion der Duplikation untersucht werden.

### 3.12 Analyse von *Bs4C*-Paralogen in *Capsicum pubescens*

#### 3.12.1 Identifikation von *Bs4C*-Paralogen in *C. pubescens*

Die Bestimmung der Volllängen- und Promotersequenz von *Bs4C-R* aus *C. pubescens* PI 235047 erfolgte mittels Amplifikation unter Verwendung mehrerer, verschiedener Oligonukleotidkombinationen (siehe A25). Die erhaltenen PCR-Produkte wurden sequenziert und die Sequenzen miteinander verglichen. Diese Sequenzanalyse ließ vermuten, dass auch in *C. pubescens* (wie in *C. annuum*) mindestens ein weiteres *Bs4C-R*-Paralog vorhanden ist (siehe A25). Mit Hilfe von spezifischen Oligonukleotiden konnte die vollständige Sequenz des *Bs4C*-Paralogs ermittelt werden (siehe A25+A26), welches nachfolgend als *CpBs4C.2-R* (Paralog aus der *C. pubescens* Akzession PI 235047) bezeichnet wird. Das entsprechende Allel aus der Akzession PI 585270 wurde ebenfalls isoliert und nachfolgend als *CpBs4C.2-S* (Paralog aus der *C. pubescens* Akzession PI 585270) benannt (siehe A26).

#### 3.12.2 Sequenzanalyse von *Bs4C*-Paralogen in *C. pubescens*

Ein Sequenzvergleich zwischen *CpBs4C.2-R* und *Bs4C-R* zeigte, dass beide *C. pubescens*-Paraloge eine Identität von 76 % zueinander aufweisen (Abb. 22, siehe A27). Vergleicht man *CpBs4C.2-R* mit den homologen Sequenzen aus *C. annuum* (CaBs4C.1 und 2), wird ersichtlich, dass CaBs4C.1 und CaBs4C.2 eine Sequenzidentität von 84 % bzw. 79 % zu *CpBs4C.2-R* zeigen (siehe A28+A29). Entfernt man jedoch die Duplikation von 26 AS aus der Sequenz von CaBs4C.2 (CaBs4C.2Del), steigt die Homologie zu *CpBs4C.2-R* auf 91 % (Abb. 22, siehe A30). Diese Vergleiche machen deutlich, dass *CpBs4C.2-R* die höchste Identität zu CaBs4C.2Del während *Bs4C-R* die höchste Identität zu CaBs4C.1 aufweist (siehe A30). Der Sequenzvergleich zwischen *CpBs4C.2-R* und *CpBs4C.2-S* ergab, dass sie sich nur in drei AS unterscheiden (siehe A26).

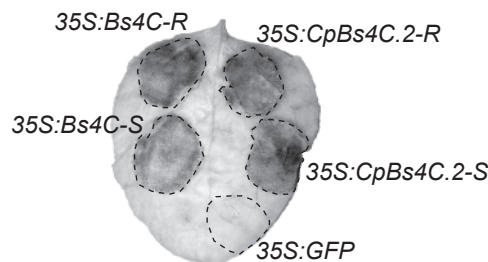
		1	2	3	4	5	6	7
CaBs4C.2	1		86.10	79.14	78.61	74.21	67.01	68.91
CaBs4C.2Del	2	26		91.93	91.30	85.98	77.38	79.64
CpBs4C.2-S	3	39	13		98.14	83.54	75.60	75.45
CpBs4C.2-R	4	40	14	3		83.54	75.60	75.45
CaBs4C.1	5	49	23	27	27		82.25	79.88
Bs4C-R	6	64	38	41	41	30		93.33
Bs4C-S	7	60	34	41	41	34	11	

**Abb. 22 Sequenzidentitäten verschiedener Bs4C-R-Homologen zueinander**

Das Alignment (siehe A30) und der Vergleich der Proteinsequenzen wurde mit dem Programm CLC erstellt. Dabei sind im oberen rechten Dreieck die Sequenzidentitäten und im unteren linken Dreieck die Anzahl von Polymorphismen angegeben.

### 3.12.3 Funktionale Analyse von *CpBs4C.2-R* und *CpBs4C.2-S*

Die Überexpressionsanalysen von *CaBs4C.2* zeigten, dass es keine HR in *N. benthamiana* induzieren kann (Stephanie Solle, FGP). Dabei liegt die Vermutung nahe, dass eine Duplikation von 26 AS die Struktur von *CaBs4C.2* und somit seine Funktionalität beeinflussen kann. *CpBs4C.2-R* und *CpBs4C.2-S* weisen diese AS-Duplikation nicht auf und könnten aus diesem Grund möglicherweise funktional sein. Diese Vermutung wurde mittels 35S-Konstrukten in *N. benthamiana* getestet. Sowohl *CpBs4C.2-R* als auch *CpBs4C.2-S* sind in der Lage, eine HR bei Überexpression zu induzieren (Abb. 23).



**Abb. 23 Die konstitutive Expression der *C. pubescens* Bs4C-Paraloge *CpBs4C.2-R* und *CpBs4C.2-S* resultiert in einer HR**

Phänotypen nach *Agrobacterium*-vermittelter Transformation von *N. benthamiana*. Die KDS der *C. pubescens* Paraloge *CpBs4C.2-R* und *CpBs4C.2-S* wurden unter Kontrolle des 35S-Promoters in einen T-DNA Vektor kloniert und *Agrobacterium* vermittelt in *N. benthamiana* exprimiert. Als Negativkontrolle diente ein 35S:GFP-Konstrukt, als Positivkontrolle 35S:Bs4C-R und 35S:Bs4C-S-Konstrukte (Strauß et al., 2012). Die Blätter wurden drei Tage nach Inokulation geerntet und zum besseren Sichtbarmachen der HR (dunkle Blattbereiche) in Ethanol entfärbt.

### 3.12.4 RT-PCR-Analysen für *CpBs4C.2-R* und *CpBs4C.2-S*

Die Ergebnisse der Überexpressionsanalyse zeigten, dass *CpBs4C.2-R* und auch *CpBs4C.2-S* funktional sind. Daraufhin stellte sich die Frage, ob auch *CpBs4C.2-R* AvrBs4-induzierbar ist. Die Analyse der *CpBs4C.2-R*-Promotersequenz konnte ein

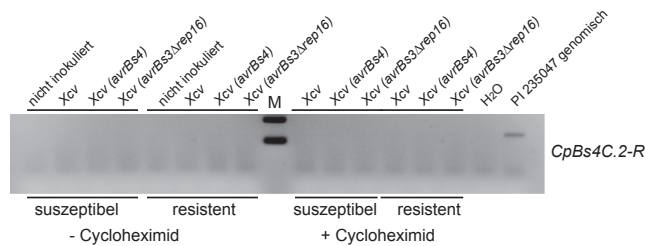
mögliches EBE 226 nt vor dem potentiellen Startkodon identifizieren (siehe A26), welches mit dem aus *CaBs4C.2* übereinstimmt (Abb. 24). Das EBE aus *CpBs4C.2-R* sowie alle in *C. annuum* identifizierten EBE-Sequenzen zeigen an Position zwei und drei C-Nukleotide (siehe 3.11.2, Abb. 24), wie das EBE von *Bs4C-S*, was vermuten lässt, dass auch *CpBs4C.2-R* nicht AvrBs4-induzierbar ist.

	Nukleotidposition im <i>EBE</i>																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
<i>CaBs4C.2</i>	T	G	C	C	A	A	A	A	A	T	A	G	C	C	C	T	C	C	A
<i>CpBs4C.2-R</i>	T	G	C	C	A	A	A	A	A	T	A	G	C	C	C	T	C	C	A
<i>CaBs4C.1</i>	T	A	C	C	A	A	A	A	A	T	A	G	C	C	C	T	C	C	C
<i>Bs4C-S</i>	T	A	C	C	A	A	A	A	A	T	A	G	T	C	C	T	C	T	C

**Abb. 24 Vergleich der EBE von *CpBs4C.2-R* mit denen von *Bs4C-S* und den *C. annuum*-Homologen *CaBs4C.1* und *CaBs4C.2*.**

Die EBE von *CpBs4C.2-R* und *CaBs4C.2* sind identisch und weisen einen Polymorphismus zum EBE von *CaBs4C.1* auf. Alle abgebildeten EBE zeigen an Position zwei und drei (rote Umrandung) C-Nukleotide, wie sie bereits bei *Bs4C-S* (nicht AvrBs4-responsiv) identifiziert wurden. Identische Nukleotide sind weiß markiert und haben einen schwarzen Hintergrund. Unterschiedliche Nukleotide sind schwarz auf weißem Hintergrund dargestellt.

Zur Analyse der Induzierbarkeit der *CpBs4C.2*-Promotoren wurden sq RT-PCR-Analysen mit paralogspezifischen Oligonukleotiden durchgeführt. Wie in Abb. 25 zu erkennen ist, konnte eine transkriptionelle Induktion von *CpBs4C.2-R* und auch *CpBs4C.2-S* durch AvrBs4 nicht nachgewiesen werden. Das verwendete cDNA-Material war identisch mit dem, welches für die sq RT-PCR-Analysen von *Bs4C-R* verwendet wurde (siehe 3.2) und auch die Zyklenzahl ist gleich (29 Zyklen). Als PCR-Kontrolle diente genomische DNA, auf der eine für *CpBs4C.2-R* spezifische Oligonukleotidkombination ein Amplifikationsprodukt lieferte (Abb. 25). Die sq RT-PCR-Analysen zeigen, dass die *CpBs4C.2*-Paraloge nicht durch AvrBs4 induziert sind, was in Übereinstimmung mit den EBE-Sequenzanalysen steht. In diesen konnte auch für *CpBs4C.2-R* die zwei C-Nukleotide nachgewiesen werden, die auch im nicht AvrBs4-responsivem EBE von *Bs4C-S* vorkommen (Abb. 24). Der Vergleich mit dem putativen EBE von *CaBs4C.2* zeigte, dass beide identisch sind (Abb. 24) und unterstützt die Vermutung, dass *CpBs4C.2-R* aus *C. pubescens* das „direkte Ortholog“ zu *CaBs4C.2* aus *C. annuum* darstellt.



**Abb. 25 Die *Bs4C*-Paraloge *CpBs4C.2-R* und *CpBs4C.2-S* sind nicht durch *AvrBs4* induziert**  
*C. pubescens*-Blätter wurden mit *Xcv*-Stämmen, die TALE-Gene exprimieren (*avrBs4* und *avrBs3Δrep16*), in An- und Abwesenheit von Cycloheximid (+/-), infiziert und 18 Stunden nach Inokulation geerntet. Für die Durchführung der sq RT-PCR-Analysen wurden paralogspezifische Oligonukleotide, verwendet. Als PCR-Kontrolle diente genomische DNA von *C. pubescens* (PI 235047), die Grundlage für eine Amplifikation mit *CpBs4C.2*-spezifischen Oligonukleotiden (390 nt) war. Die erhaltenen PCR-Produkte wurden auf einem 1 %igen Agarosegel aufgetrennt und mit Ethidiumbromid angefärbt. M = GeneRuler 1kb DNA Größenstandard

### 3.13 Identifikation von *Bs4C-R*-Homologen in Solanaceae-Spezies

#### 3.13.1 Phänotypische Analyse

Um zu untersuchen, ob auch in anderen Gattungen der Familie der Solanaceen *Bs4C*-Homologe vorhanden sind, wurden 36 verschiedene *Capsicum*-Spezies sowie sieben weitere Spezies aus der Familie der Solanaceen analysiert. Eine detaillierte Auflistung aller getesteten Pflanzen befindet sich im Anhang (siehe A32). Die Pflanzen wurden zunächst mit dem WT *Xanthomonas* Stamm (*Xcv* 85-10) sowie *AvrBs4*-exprimierenden Xanthomonaden inokuliert. Drei bis fünf Tage nach Inokulation wurden die Pflanzen hinsichtlich ihrer phänotypischen Reaktionen untersucht. Es wurden wenigstens vier Pflanzen pro Spezies infiltriert, die bis auf zwei Ausnahmen (siehe Anhang, *C. pubescens* CAP1480 und *C. eximium* CGN21502), die gleichen *AvrBs4*-Reaktionen auf allen Pflanzen zeigten. Von den sieben *Solanaceae*-Arten, die keine *Capsicum*-Spezies sind, wies keine eine HR nach Inokulation mit *AvrBs4*-exprimierenden Xanthomonaden auf. Dies lässt darauf schließen, dass diese Spezies kein oder ein nicht-funktionales *Bs4C-R*-Homolog besitzen. Alle Infiltrationsbereiche der zwei *C. chacoense*-Arten zeigten Zelltodreaktionen, die durch *Xanthomonas* (allein) hervorgerufen wurden. Bei 17 der 36 *Capsicum*-Arten konnte eine *AvrBs4*-induzierte HR beobachtet werden, so auf 13 *C. pubescens*-Arten, einer *C. cardenasii*, einer *C. eximium* und zwei *C. chacoense* Spezies. Die Spezies *C. chinense* müsste nochmal nachgetestet werden, da die *AvrBs4*-Reaktionen nicht eindeutig waren.

#### 3.13.2 Sequenzanalyse und Vergleich der EBE

Um potentiell vorhandene *Bs4C-R*-Homologe zu identifizieren, wurde von allen getesteten Spezies Blattmaterial geerntet und genomische DNA isoliert. Mit



verschiedenen *Bs4C-R*-spezifischen Oligonukleotidkombinationen (Promoter- und KDS) wurde versucht die *Bs4C*-Orthologen zu amplifizieren und zu sequenzieren. Ein Überblick über alle Sequenzinformationen ist im Anhang (A31+A32) dargestellt. Leider war kein vollständiger Datensatz verfügbar und somit konnten nachfolgend nur begrenzt Auswertungen vorgenommen werden.

Es konnten die KDS von *Bs4C-R*-Homologen in 28 und potentielle EBEs in den Promotersequenzen von 17 *Capsicum*-Spezies ermittelt werden (A31+A32). Die identifizierten EBE konnten in fünf verschiedene Klassen eingeteilt werden (Tab. 4).

- Zur Klasse 1 gehören Promotoren, welche das  $EBE_{AvrBs4Bs4C-R}$  aufweisen und durch AvrBs4 induzierbar sind (siehe 3.6). In Übereinstimmung damit zeigten alle sieben Spezies bei der phänotypischen Analyse mit AvrBs4-exprimierenden Xanthomonaden eine HR (Tab. 4, siehe A32).
- Die Klasse 2 ( $EBE_{AvrBs4Bs4C-R/S}$ ) stellt eine Mischung aus den  $EBE_{AvrBs4Bs4C-R}$  und  $EBE_{AvrBs4Bs4C-S}$  dar. Sie besitzt ein T-Nukleotid an Position zwei, wie das  $EBE_{AvrBs4Bs4C-R}$ , und ein C-Nukleotid an Position drei, wie das  $EBE_{AvrBs4Bs4C-S}$  (Tab. 4). Die phänotypischen Analysen dieser Pflanzen nach Infektion mit AvrBs4-exprimierenden Xanthomonaden ergaben keine einheitlichen Ergebnisse. Zwei *C. pubescens*-Spezies zeigten eine HR, zwei andere nicht. EMSA-Analysen, welche von Christina Wolf durchgeführt wurden, von AvrBs4 mit dem  $EBE_{AvrBs4Bs4C-R}$  bzw. dem  $EBE_{AvrBs4Bs4C-R/S}$  zeigten, dass die Affinität der Misch-Box geringer ist als die des  $EBE_{AvrBs4Bs4C-R}$  (Christina Wolf, unveröffentlichte Daten).
- Die Klasse 3 beinhaltet das putative  $EBE_{AvrBs4Bs4C-S}$ , welches in drei Spezies gefunden wurde, zu denen zwei *C. pubescens*- und eine *C. tovarii*-Spezie gehörten. Die Expression der *Bs4C*-Allele dieser 3 Spezies können nicht durch AvrBs4 induziert werden (siehe 3.6, Tab. 4, siehe A32).
- Die Klasse 4-zugehörige Spezies wiesen das putative  $EBE_{AvrBs4CaBs4C.I}$  auf, welches in *C. annuum* ECW und cv. CM334 identifiziert wurde. Dies ist nicht durch AvrBs4 induzierbar wie bereits gezeigt wurde (Tab. 4, Abb. 20). Bei den Spezies, die der Klasse 4 und 5 zugeordnet wurden, konnte ebenfalls keine HR nach Infektion mit AvrBs4-exprimierenden Xanthomonaden beobachtet werden (siehe A32).

- Die Klasse 5 ( $EBE_{AvrBs4}CbBs4C.1$ , *Cb* - *C. baccatum*) konnte in zwei *C. baccatum*-Spezies identifiziert werden. Sie leitet sich vom putativen  $EBE_{AvrBs4}CaBs4C.1$  ab und weist einen zusätzlichen Polymorphismus an letzter Stelle auf und ist mit hoher Wahrscheinlichkeit ebenfalls nicht durch AvrBs4 induzierbar (Tab. 4).

**Tab. 4 Übersicht der EBEs in den analysierten *Capsicum*-Spezies**

EBE-Klasse	Bezeichnung des EBE	Sequenz des EBE	AvrBs4-Reaktion <sup>a)</sup>	Spezies
1	$EBE_{AvrBs4}Bs4C-R$	TATAAAAAATAGTCCTCTC	HR	<i>C. pubescens</i> <sup>1)</sup>
2	$EBE_{AvrBs4}Bs4C-R/Bs4C-S$	TATCAAAAAATAGTCCTCTC	HR/keine HR	<i>C. pubescens</i> <sup>2)</sup>
3	$EBE_{AvrBs4}Bs4C-S$	TACCAAAAAATAGTCCTCTC	keine HR	<i>C. pubescens</i> , <i>C. tovarii</i> <sup>3)</sup>
4	$EBE_{AvrBs4}CaBs4C.1$	TACCAAAAAATAGCCCTCC	keine HR	<i>C. eximium</i> <sup>4)</sup>
5	$EBE_{AvrBs4}CbBs4C.1$	TACCAAAAAATAGCCCTCCA	keine HR	<i>C. baccatum</i> <sup>5)</sup>

<sup>1)</sup> *C. pubescens* CAP217, CAP1694, CAP867, CAP907, CAP1481, CAP1482; CAP1483  
<sup>2)</sup> *C. pubescens* CAP468, CAP273, CAP870, CGN19243  
<sup>3)</sup> *C. pubescens* CAP1486, CAP1492; *C. tovarii* Eshba CGN22876  
<sup>4)</sup> *C. eximium* CAP1035/94  
<sup>5)</sup> *C. baccatum* CAP874, CAP1475  
<sup>a)</sup> Infiltration mit AvrBs4-exprimierenden *Xanthomonaden*, HR - hypersensitive Reaktion

### 3.13.3 Sequenzanalyse und Vergleich der *Bs4C-R*-homologen Sequenzen

Von den 36 untersuchten Spezies wurde in 17 durch AvrBs4 eine HR induziert (siehe A32). Für 15 dieser 17 Spezies konnte die KDS des *Bs4C-R*-Allel bestimmt werden (siehe A31). Die Sequenz von 11 Allelen war identisch mit der von *Bs4C-R*. Vier Spezies wiesen zwischen einem und bis zu 18 AS-Polymorphismen in der *Bs4C-R*-homologen Sequenz auf und zeigten somit eine Sequenzidentität von (minimal) 89 % zu *Bs4C-R* (A31).

In den 19 Spezies, die keine AvrBs4-abhängige HR zeigten (siehe A32), konnte die Sequenz von 13 Allelen bestimmt werden. Dreimal konnte das *Bs4C-S*-Allel identifiziert werden, während zehn Allele Polymorphismen in unterschiedlicher Anzahl gegenüber dem *Bs4C-R*-Allel aufwiesen. Die geringste Sequenzidentität (85 %) zeigte ein Allel, bei welchem 24 Polymorphismen verglichen mit *Bs4C-R* festgestellt wurden (siehe A31).

In keiner der Proteinsequenzen der insgesamt 28 analysierten Spezies, für die das *Bs4C*-Allel bestimmt wurde, konnte ein vorzeitiges Stoppkodon identifiziert werden (siehe A31). Alle Sequenzen bzw. eine Darstellung der identifizierten Polymorphismen in den einzelnen Allelen ist im Anhang (A31+A32) zu finden.

## 4 Diskussion

### 4.1 RNA-Seq kann zur Isolierung von TALE-induzierten Resistenzgenen verwendet werden

Im Rahmen dieser Arbeit konnte mit Hilfe von NGS und Transkriptomanalysen das Resistenzgen *Bs4C-R* aus Paprika isoliert werden. Im Vergleich zum klassischen kartengestützten (*map-based cloning*)-Ansatz (Peters et al., 2003), der bislang für die Isolierung zahlreicher *R*-Gene genutzt wurde (Gu et al., 2008; Iyer and McCouch, 2004; Römer et al., 2007; Schornack et al., 2004; Tian et al., 2014), ist der NGS-basierte Ansatz schnell und einfach. So dauerte beispielsweise die kartengestützte Isolierung des *R*-Gens *Bs3* aus Paprika 13 Jahre, während *Bs4C-R* in vier Jahren isoliert werden konnte (Jordan et al., 2006; Pierre et al., 2000; Römer et al., 2007; Strauß et al., 2012). Für die Isolierung von *Bs3* war die Etablierung einer Kartierungspopulation (>4000 Pflanzen), die Identifizierung und Verdichtung von *Bs3*-gekoppelten molekularen Markern (47 Marker) sowie die Erstellung einer BAC-Bibliothek und *in planta* Komplementationsanalysen notwendig (Jordan et al., 2006; Römer et al., 2007). Im Vergleich dazu gestaltete sich die Isolierung von *Bs4C-R* weniger zeit- und arbeitsintensiv. Es wurde lediglich *Xanthomonas*- (+/- AvrBs4) infiziertes Paprikamaterial benötigt, mit welchem RNA-Seq-Analysen durchgeführt wurden. Da weder eine Kartierungspopulation noch eine BAC- oder YAC-Bibliothek vorhanden sein muss und auch die aufwendige Identifizierung von molekularen Markern entfällt, ist dieser Ansatz zur Isolierung von Resistenzgenen aus allen wichtigen Kulturpflanzen geeignet. Jedoch kann der Ansatz nur für *R*-Gene verwendet werden, die durch TAL-Effektoren aus *Xanthomonas* transkriptionell induziert werden. Das heißt, er eignet sich nicht für *R*-Gene, bei denen die Erkennung von TAL-Effektoren durch direkte oder indirekte Interaktion zwischen R-Protein und TAL-Effektor vermittelt wird. Ein Beispiel dafür ist *Bs4S* aus Tomate (*Solanum lycopersicum*), welches für ein R-Protein der NBS-LRR-Klasse kodiert (Schornack, 2006; Schornack et al., 2004).

Um zu testen, ob die Erkennung auf transkriptioneller Induktion oder Interaktion beruht, können kurzerhand Analysen mit TAL-Effektormutanten, welche Deletionen in AD und NLS aufweisen, durchgeführt werden. Allerdings stellt *Bs4S* eine

Ausnahme dar, da es bislang das einzige TAL-erkennende *R*-Gen ist, das auch AvrBs4-Deletionsderivate erkennt, die nur noch aus 3,5 repeats bestehen und vermutlich direkt oder indirekt mit AvrBs4 bzw. mit TAL-Effektoren interagiert, während Bs4C-R nur das AvrBs4-Volllängenprotein „erkennt“ (siehe 1.6) (Schornack et al., 2004; Schornack et al., 2005; Strauß et al., 2012). Die anderen, bislang isolierten *R*-Gene, wie *Bs3* aus Paprika oder *Xa7*, *Xa10* und *Xa27* aus Reis, die „Erkennung“ von TAL-Effektoren vermitteln, werden durch die jeweiligen TALEs auch transkriptionell induziert (Abb. 4) (Chen et al., 2008; Gu et al., 2008; Gu et al., 2005; Römer et al., 2007; Tian et al., 2014). Demnach sollte der RNA-Seq-Ansatz für die meisten TAL-*R*-Gene anwendbar sein und aufgrund des deutlich geringeren Arbeitsaufwandes und Materialverbrauchs unkompliziert umsetzbar sein.

Interessanterweise besitzen zahlreiche Stämme des Phytopathogens *Ralstonia solanacearum* TAL-ähnliche Effektoren, sogenannte RipTALs (*Ralstonia injected proteins*), für die kürzlich gezeigt wurde, dass sie ebenso wie *Xanthomonas*-TALEs, als eukaryotische Transkriptionsfaktoren fungieren und die Expression von Zielgenen induzieren (de Lange et al., 2013). Eine Sichtung wichtiger Kulturpflanzen bzw. Wildspezies mit RipTALs und deren AD- und NLS-Derivaten könnte es ermöglichen TAL-induzierte *R*-Gene zu identifizieren, welche mittels NGS- und transcriptional profiling-Ansatz isoliert werden können. Ein weiterer Vorteil dieses Ansatzes besteht darin, dass Sichtungen nur mit dem TAL-Protein vorgenommen werden können, welche bspw. mittels *Pseudomonas fluorescens* transloziert werden anstatt mit dem zur S2-Sicherheitsklasse-gehörigem Pathogen selbst und dadurch der sicherheitstechnische Aufwand wesentlich geringer ist (Upadhyaya et al., 2014). *Ralstonia* rangiert auf Platz 2 der Liste der 10 bedeutendsten bakteriellen Pathogene (Mansfield et al., 2012) und zeichnet sich durch seine weite geographische Verbreitung sowie sein enorm breites Wirtsspektrum aus. Bislang konnten jedoch bis auf das *A. thaliana* R-Protein RRS1-R, das den *Ralstonia* GMI1000 Effektor PopP2 erkennt, kaum monogenische Resistenzen gegen *Ralstonia* isoliert werden (Deslandes et al., 2002).

## 4.2 Kritische Parameter für RNA-Seq-basierte Isolierung von TAL-induzierten *R*-Genen

### 4.2.1 Auswahl des Zeitpunktes der Probennahme

Die Methode des transcriptional profiling zur Isolierung von *Bs4C-R* hat sich als effizient erwiesen. Jedoch hat sich herausgestellt, dass es einige Punkte gibt, die es zu beachten gilt für die Isolierung von TAL-*R*-Genen mittels NGS-basierten Transkriptomanalysen.

Für die Isolierung des *R*-Gens *Bs4C-R* aus *C. pubescens* wurden zwei RNA-Seq-Ansätze, eine 454- und eine Illumina-basierte Transkriptanalyse, initiiert (siehe 3.1, 3.8). Jedoch konnte *Bs4C-R* erst im zweiten Ansatz isoliert werden (siehe 3.1). Der Grund hierfür lag in den unterschiedlichen Erntezeitpunkten des *Xanthomonas*- (+/- AvrBs4) infizierten Paprikamaterials. Für den ersten Ansatz wurde nur zu einem Zeitpunkt und zwar 12 Stunden nach Inokulation geerntet, während die Ernte für den zweiten Ansatz alle 6 Stunden in einem Zeitfenster von 0 bis 24 Stunden erfolgte (siehe 3.1, 3.2). Die Überlegungen waren, dass mit der Wahl des frühen Erntezeitpunktes im ersten Ansatz die Zahl der identifizierten Transkripte von Sekundärgenen, welche Zelltod-assoziierte Gene, indirekt-induzierte Gene, und PR-Gene (*pathogenesis-related*) mit einschließen, so gering wie möglich gehalten werden sollte.

Unsere Illumina-basierten Transkriptanalysen machten aber deutlich, dass *Bs4C-R*-Transkript erst in der 18 Stunden-Probe auftrat (Tab. 1). Weitere RT-PCR-Analysen zeigten außerdem, dass die *Bs4C-R* Transkriptinduktion nicht vor 15 Stunden nachweisbar ist (siehe 3.2). Für das kürzlich isolierte TAL-induzierte *R*-Gen *Xa10* und für *Bs3* waren erste Transkripte bereits nach 6 bzw. 9 Stunden nachweisbar, was zeigt, dass diese drei *R*-Gene sich hinsichtlich ihres Expressionszeitpunktes unterscheiden (Römer et al., 2007; Tian et al., 2014). Daraus kann geschlussfolgert werden, dass es nicht möglich ist, eine generelle Aussage über die Induktion von TAL-induzierten *R*-Gene zu treffen. Somit sollten für die Isolierung weiterer *R*-Gene stets mehrere Erntezeitpunkte gewählt werden.

### 4.2.2 Erstellung der cDNA-Bibliothek

Für die Erstellung von cDNA-Bibliotheken können entweder random oder oligo dT-Oligonukleotide genutzt werden. Da zu Beginn der Arbeiten keine Genomsequenz von *C. pubescens* vorlag, ermöglichte die Verwendung von oligo dT-

Oligonukleotiden die eindeutige Zuordnung der Sequenzreads zueinander, die in contigs zusammengefasst wurden (siehe 2.14). Folglich ließ die Anzahl, der zu einem contig gehörigen Sequenzreads, Rückschlüsse über dessen Expressionslevel zu. Da für die TAL-induzierten *R*-Gene *Bs3*, *Xa10* und *Xa27* gezeigt wurde, dass ihre Expression streng reguliert ist und Transkript für *Bs3* in Abwesenheit des TAL-Effektors AvrBs3 nicht nachgewiesen konnte, wurde angenommen, dass dies auch für *Bs4C* und AvrBs4 zutrifft. Das heißt, die Expression von *Bs4C* in Abwesenheit von AvrBs4 ist so gering, dass sie unterhalb der Detektionsgrenze für Real-Time und Reverse Transkriptions-PCR Analysen liegt. In Anwesenheit von AvrBs4 ist die *Bs4C*-Expression jedoch stark induziert ist (Gu et al., 2005; Römer et al., 2007; Schornack et al., 2013; Tian et al., 2014).

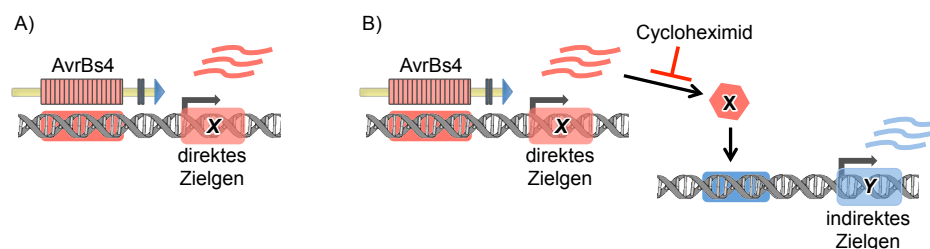
Kürzlich wurde eine weitere Methode (NanoCAGE kombiniert mit CAGEscan) entwickelt, die es erlaubt 5'-Enden von cDNAs anzureichern (Plessy et al., 2010; Salimullah et al., 2011). Der Vorteil gegenüber der für *Bs4C*-verwendeten Methode besteht darin, dass Informationen über die Transkriptionsstartpunkte der induzierten Gene erhalten werden. Damit ist eine schnellere und zuverlässigere Bestimmung der EBE möglich, da diese in einem Bereich von 40 bis 100 Nukleotiden *upstream* zum Transkriptionsstart liegen sollten (Kay et al., 2009; Römer et al., 2009a). Außerdem sollte es eine effizientere Identifizierung von direkten Zielgenen und des *R*-Gens erlauben, da nur diese ein EBE aufweisen sollten. Da im Falle der *Bs4C*-Isolierung kein CHX (siehe 4.2.3) bei der Erstellung der Transkriptbibliotheken verwendet wurde, umfassten sequenzierte reads sowohl direkte als auch indirekte Zielgene von AvrBs4, was in der großen Anzahl an potentiellen *Bs4C*-Kandidaten und deren erneuter Analyse mit CHX-behandeltem Material resultierte (siehe 3.8).

Des Weiteren sind 5'-Bibliotheken vor allem auch dann von Vorteil, wenn die Genomsequenz nicht verfügbar ist. Vermutlich wird der Aufwand, mittels PCR-Walking (Leoni et al., 2011) Promotoren zu identifizieren und in diesen nach EBE zu „suchen“, geringer sein, da weniger PCR-Schritte notwendig sind und weniger Probleme mit Artefakten aus Genfamilien auftreten werden. Die Isolierung von *Bs4C-R* unter Verwendung des oligo dT-Ansatzes verlief sehr gut und war erfolgreich. Aber hier stand, wenn auch nur partiell, die Genomsequenz einer verwandten Spezies (*C. annuum* cv. CM334) (Kim et al., 2014) zur Verfügung, was eine relative zuverlässige Vorhersage von ORF und Promotoren der analysierten Gene zuließ.

Jedoch wurden für *Bs4C-R* auch PCR-Walking-Analysen (Leoni et al., 2011) durchgeführt um das EBE im Promoter zu identifizieren und seine Sequenz eindeutig zu bestimmen. Für die Isolierung weiterer TAL-induzierter *R*-Gene kann sowohl unser Ansatz als auch die Methode des NanoCAGE kombiniert mit CAGEscan verwendet werden, da für viele Spezies bereits Genomsequenzen bzw. Sequenzen verwandter Spezies zur Verfügung stehen. Außerdem würde die Verwendung von CHX direkte Zielgene identifizieren, in denen mit Hilfe der verfügbaren Genomsequenzen EBE bestimmt werden und somit das *R*-Gen isoliert werden kann.

#### 4.2.3 Nutzung von Cycloheximid zur Identifizierung des *R*-Gens als direktes Zielgen unter der Gesamtheit der TAL-induzierten Gene

Für das TAL-induzierte *R*-Gen *Bs4C-R* konnte gezeigt werden, dass es ein direktes Zielgen von AvrBs4 ist und ein EBE im Promoter aufweist. Eine Differenzierung zwischen direkten und indirekten Zielgenen war aber allein durch die RNA-Seq-basierten Transkriptanalysen nicht möglich. Daher wurden anschließend sq RT-PCR-Analysen in An- und Abwesenheit von CHX durchgeführt, um einmal die Daten der Transkriptanalysen zu bestätigen und um vor allem zwischen direkten (mit EBE<sub>AvrBs4</sub>) und indirekten (ohne EBE<sub>AvrBs4</sub>) AvrBs4-Zielgenen unterscheiden zu können (siehe 3.8, Abb. 16, Abb. 26).



**Abb. 26 Unterscheidung zwischen direkten (A) und indirekten (B) AvrBs4-Zielgenen durch Cycloheximid.**

A) Bei direkten AvrBs4-Zielgenen (X) erfolgt die Bindung von AvrBs4 an ein EBE im Promoter und anschließend die Induktion der Expression. B) Die Transkription von indirekten Zielgenen (Y) wird durch Proteine direkter/primärer AvrBs4-Zielgenen (X) induziert. Indirekte Zielgene weisen kein EBE im Promoter auf. CHX inhibiert die Proteinneusynthese und somit die Bildung von direkten AvrBs4-Zielgenprodukten (X). Dadurch können indirekten Zielgene (Y) nicht induziert werden.

Direkte Zielgene weisen ein EBE im Promoter auf, das der TAL-Effektor nach Bindung induziert. Indirekte Zielgene können durch Proteine direkter AvrBs4-Zielgene induziert werden, weisen aber kein EBE im Promoter auf. Cycloheximid ist ein Antibiotikum, welches die Proteinneusynthese inhibiert und somit die Bildung von direkten AvrBs4-Zielgenprodukten verhindert. Dadurch wird die Transkription von indirekten Genen unterbunden und eine Unterscheidung zwischen direkten und

indirekten Zielgenen ermöglicht (Abb. 26). Für andere TAL-Effektoren konnte mit Hilfe von CHX bereits erfolgreich zwischen direkten und indirekten Zielgenen differenziert werden (Hu et al., 2014; Kay et al., 2007; Römer et al., 2007).

Für 26 Kandidaten unserer Transkript-Analysen für AvrBs4 wurden sq RT-PCR-Analysen mit CHX durchgeführt und für 10 Kandidaten lieferten sie eindeutige Ergebnisse. Dies ließ eine Einteilung in direkte (5) und indirekte (5) AvrBs4-Zielgene zu (siehe 3.8, Tab. 5) und zeigte, dass CHX zur Differenzierung sehr gut geeignet ist. Daher sollte generell für die RNA-Seq-basierte Isolierung von TAL-*R*-Genen CHX-behandeltes Material verwendet werden. In Abhängigkeit von der Gesamtanzahl der Proben, die durch die Wahl mehrerer Zeitpunkte (siehe 4.2.1), Infektionen in resistenter und susceptibler Akzession sowie biologischer Replikate bestimmt wird, können sich RNA-Seq-Analysen kostenintensiv gestalten, wenn zusätzlich CHX-behandeltes Material in die Analysen einbezogen wird. Die Beschränkung auf ausschließlich CHX-behandeltes Material würde die Kosten erheblich senken und sollte nur direkte Zielgene (darunter auch das *R*-Gen) identifizieren, womit aufwendige RT-PCR-Analysen im Nachhinein zur Differenzierung von direkten und indirekten Zielgenen entfallen würden. Allerdings war für 11 Gene nach den RT-PCR-Analysen in Anwesenheit von CHX die Differenzierung nicht möglich, da sowohl in der resistenten als auch/oder in der susceptiblen Akzession, ungeachtet der Infektion (*Xcv* +/- TALE), Transkript des jeweiligen Gens nachweisbar war. Demgegenüber konnte in den Proben ohne CHX-Behandlung nur die spezifische Induktion des Kandidatengens durch AvrBs4 und nur in der resistenten Akzession detektiert werden (siehe 3.8, Tab. 5). Aus der Literatur ist bereits bekannt, dass CHX nicht nur die Proteinbiosynthese inhibiert sondern auch in der Lage ist, Gene zu induzieren sowohl in Säugetieren als auch in Pflanzen (Almendral et al., 1988; Horvath et al., 1998; Zipfel et al., 1989). Dies konnte anhand der sogenannten SAUR-Gene (*small auxin up RNAs*) aus Soja und der Hitzeschockproteine aus Reis gezeigt werden (Agarwal et al., 2011; Franco et al., 1990). Ihre Induktion ist bei Verwendung von CHX um ein Vielfaches stärker. Diese Beispiele und die Ergebnisse der RT-PCR-Analysen in Anwesenheit von CHX zeigen, dass die Gefahr besteht, Falschpositive zu isolieren, wenn man CHX verwendet. Allerdings würden diese Kandidaten auch nicht in RNA-Seq-Analysen mit ausschließlich CHX-behandelten Proben identifiziert werden können, da Transkript dieser Gene in allen Proben detektiert werden würde und diese Gene somit nicht mehr differentiell induziert wären.



Jüngst wurde gezeigt, dass CHX die Expression eines TAL-S-Gens reprimiert (Cernadas et al., 2014) und somit die Gefahr besteht, dass potentielle Kandidatengene ausgeschlossen werden könnten. Demzufolge wäre es am sinnvollsten RNA-Seq-Analysen mit CHX-behandeltem und -unbehandeltem Material durchzuführen. Der Vergleich der Transkriptzahlen durch bioinformatische Analysen sollte schnell ermöglichen, *R*-Kandidaten aus den Daten zu extrahieren und zwischen direkten und indirekten Kandidaten zu verifizieren. Für diese wird erwartet, dass sie nur in der resistenten Akzession ausschließlich durch den TALE induziert und zusätzlich CHX-insensitiv sind.

Somit sollten, trotz der Vielzahl an Proben und der damit verbundenen Kosten, sowohl mehrere Erntezeitpunkte des induzierten und nicht-induzierten Materials der suszeptiblen und resistenten Akzession als auch CHX-behandelte Proben in die RNA-Seq-Analysen eingeschlossen werden. Damit werden material- und zeitaufwendige Analysen im Nachhinein vermieden.

#### 4.2.4 Vorhersageprogramme allein genügen nicht zur Isolierung von TALE-induzierten Resistenzgenen

Nach der Entdeckung des TALE-Codes wurden Vorhersageprogramme entwickelt, die es ermöglichen, *in silico* EBE für TAL-Effektoren in Promoter-Sequenzen vorauszusagen. Derzeit sind verschiedene Programme verfügbar, deren Vorhersage auf der Analyse von experimentell validierten TALE-EBE-Interaktionen basiert und die neben der RVD-Spezifität (Storyteller, Target Finder und Talvez) auch die Affinität der einzelnen RVDs (Talgetter) berücksichtigen (Doyle et al., 2012; Grau et al., 2013; Perez-Quintero et al., 2013). Diese Programme wurden auch schon erfolgreich genutzt, um in Genomen potentielle Zielgene für TAL-Effektoren vorherzusagen (Booher and Bogdanove, 2014; Cernadas et al., 2014; Noel et al., 2013).

Demnach könnte es auch möglich sein, *in silico* *R*-Gene anhand ihrer EBE zu identifizieren. Voraussetzung dafür ist die Verfügbarkeit der Genomsequenz des resistenten Genotyps und idealerweise auch die des suszeptiblen Genotyps, in denen man mit Hilfe der Programme TALE-Zielgene vorhersagen lassen kann. Innerhalb dieser potentiellen TALE-Zielgene würde man dann A) nach Genen suchen, die sowohl im resistenten als auch im suszeptiblen Genotyp präsent sind und Polymorphismen in ihren EBE aufweisen.

Tab. 5 Vergleich der Daten der RNA-Seq- und sq RT-PCR-Analysen

Kandidaten	Transkriptanzahl RNA-Seq-Analysen <sup>1)</sup>						Ergebnisse der sq RT-PCR-Analysen		
	S	S- 24 hpi	S+ 24 hpi	R	R- 24 hpi	R+ 24 hpi	(-CHX) induziert		+CHX
							in	durch	Zielgen
Candidate2	0	0	0	0	1	157	R	<i>Xcv(avrBs4)</i>	C*
Candidate3	0	0	0	0	1	128	R	<i>Xcv(avrBs4)</i>	C*
Cluster1923	0	0	0	0	23	456	R	<i>Xcv(avrBs4)</i>	C*
Cluster12600	0	0	0	0	0	87	R	<i>Xcv(avrBs4)</i>	direkt
Candidate4	2	4	15	3	3	397	R	<i>Xcv(avrBs4)</i>	C*
Candidate5	0	10	27	0	0	301	R	<i>Xcv(avrBs4)</i>	indirekt
Candidate6	0	17	28	1	0	322	R	<i>Xcv(avrBs4)</i>	indirekt
Candidate7	0	6	11	0	0	110	R	<i>Xcv(avrBs4)</i>	C*
Candidate9	3	10	7	8	1	125	R	<i>Xcv(avrBs4)</i>	C*
Candidate10	0	19	17	0	2	211	R	<i>Xcv(avrBs4)</i>	C*
Candidate11	0	15	24	0	1	156	R	<i>Xcv(avrBs4)</i>	C*
Cluster13224	0	4	7	0	0	77	R	<i>Xcv(avrBs4)</i>	C*
Contig1919	3	9	17	3	6	131	R	<i>Xcv(avrBs4)</i>	C*
Cluster8984	0	7	6	0	2	86	R	<i>Xcv(avrBs4)</i>	indirekt
Cluster11620	2	6	4	0	1	60	R	<i>Xcv(avrBs4)</i>	indirekt
Cluster699	0	60	49	0	30	1006	R	<i>Xcv(avrBs4)</i>	C*
Candidate1	0	0	0	0	4	611	R+S	<i>Xcv</i>	n.g.
Contig1545	0	0	0	0	8	437	R+S	<i>Xcv</i>	n.g.
Cluster15502	0	0	18	1	1	11	R+S	<i>Xcv(avrBs4)</i>	direkt
Cluster8248	0	0	53	0	0	27	R+S	<i>Xcv(avrBs4)</i>	direkt
Contig630	4	7	244	6	7	132	R	<i>Xcv(avrBs4)</i>	direkt
Candidate15	0	109	91	0	1	198	R+S	<i>Xcv(avrBs4)</i>	indirekt
Cluster14681	1	1	6	1	1	15	R+S	<i>Xcv</i>	n.g.
Candidate14	0	53	51	0	1	113	R+S	<i>Xcv</i>	n.g.
Candidate16	0	123	158	0	0	141	R+S	<i>Xcv</i>	n.g.
Cluster9054	0	4	59	0	0	1	S	<i>Xcv(avrBs4)</i>	direkt

S – susceptible *C. pubescens*-Akzession PI585270  
 R – resistente *C. pubescens*-Akzession PI235047  
 - Pflanzenmaterial mit *Xcv* (WT) infiziert  
 + Pflanzenmaterial mit *avrBs4*-exprimierenden *Xcv* infiziert  
*Xcv* – *Xanthomonas campestris* pv. *vesicatoria*  
*Xcv(avrBs4)* – *avrBs4*-exprimierender *Xanthomonas campestris* pv. *vesicatoria*-Stamm  
 n.g. – nicht getestet  
 C\* - einige Cycloheximid-behandelten Proben zeigten höhere Transkriptmengen verglichen zu unbehandelten Proben  
<sup>1)</sup> Die Zahlen geben die Anzahl der detektierten Transkripte für den jeweiligen Kandidaten wieder.

Es wäre aber auch möglich B) Gene zu identifizieren, die im resistenten aber nicht im suszeptiblen Genotyp „auftauchen“. Jedoch wurde für alle bekannten TAL-induzierten *R*-Gene gezeigt, dass susceptible Allele (*Bs3-E*, *xa27* und *Bs4C-S*) relevante Polymorphismen in ihren EBE aufweisen und somit die Möglichkeit besteht, dass die EBE nicht mehr vorhergesagt werden sowie es für das EBE im *Bs4C-S*-Promoter der Fall war (Römer et al., 2007; Römer et al., 2010b; Römer et al., 2009b; Strauß et al., 2012).

Mittels Targetfinder wurden im Reispromoterom pro TALE bis zu 600 potentielle Zielgene vorhergesagt, was eine zu große Zahl darstellt um sie mittels Reportersystem (RT-PCR-Analysen) zu analysieren (Cernadas et al., 2014). Und auch wenn die Zahl der Kandidaten durch stringenter Vorhersagen (Veränderung der

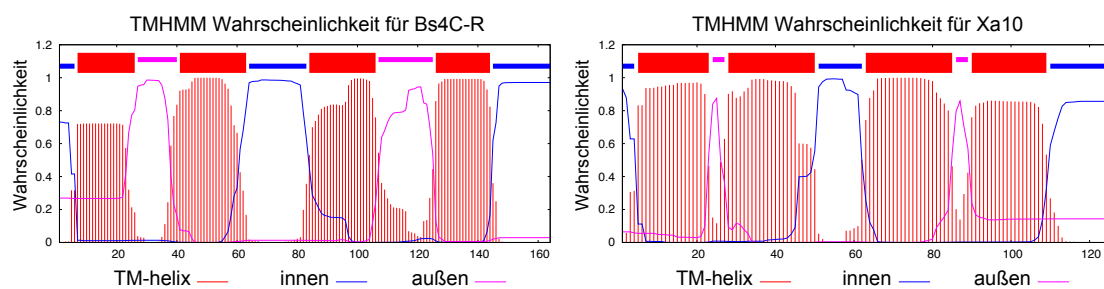
Vorhersageparameter) verringert werden kann, so sind es vermutlich trotzdem zu viele, um sie alle transient mittels GUS-Analysen (EBE) oder Überexpressions-Analysen (KDS) zu testen. Zudem konnte gezeigt werden, dass nicht alle identifizierte EBEs durch TALEs aktiviert werden können, was durch mögliche epigenetische Modifikationen bedingt sein kann (Bultmann et al., 2012; Cernadas et al., 2014; Deng et al., 2012). Das heißt, man erhält durch die Vorhersageprogramme eine große Zahl an „falsch positiven“ Gene, die man mit analysieren würde. Der Vergleich der vorhersagten Gene mit „realen“ Transkriptdaten machte aber deutlich, dass der Überlapp sehr gering ist, und die Zahl der induzierten Gene für die TALEs zwischen 0 und 7 lag (Cernadas et al., 2014). Am Beispiel von *Bs4C-R* zeigte sich ebenfalls die Limitation der Vorhersageprogramme, da das EBE im *Bs4C-R*-Promoter nicht vorhergesagt werden konnte, obwohl unsere Analysen zeigten, dass es eindeutig durch AvrBs4 induziert ist (siehe 3.6) (Strauß et al., 2012). Das EBE<sub>AvrBs4</sub>*Bs4C-R* weist insgesamt 5 Unterschiede zu einem „perfekten“ (anhand des TAL-Code vorhergesagten) EBE von AvrBs4 auf und wurde durch ein Alignment (ClustalW) eines „perfekten“ EBE auf die *Bs4C-R*-Promotersequenz identifiziert und experimentell mittels GUS-Analysen validiert (siehe 3.6). In der Tat findet man allgemein „nicht-perfekte“ EBE in TAL-induzierten Genen (Römer et al., 2007; Römer et al., 2010a; Tian et al., 2014), die unter Umständen von den Programmen nicht mehr als EBE „erkannt“ werden. Daher ist es am sinnvollsten TALE-Code-basierte Vorhersageprogramme stets mit Transkriptanalysen (z.B. RNA-Seq) zu kombinieren, da dies die Identifizierung TALE-induzierter *R*-Gene vereinfacht und effizient gestaltet.

#### **4.3 *Bs4C-R* gehört zur Gruppe der TALE-induzierten Resistenzgene und kodiert für ein putatives Transmembranprotein**

Für resistente Pflanzen wurde gezeigt, dass TALEs an EBE in *R*-Gen-Promotoren binden (siehe 1.6, Abb. 4). Die „Erkennung“ der TAL-Effektoren vermittelt hier der Promoter, die *R*-Proteine fungieren indes als Exekutoren und lösen eine Resistenzreaktion aus, die häufig mit einer Zelltodreaktion (HR) des infizierten Gewebes, einhergeht (siehe 1.6). Interessanterweise zeigen die bislang identifizierten TAL-*R*-Proteine auf AS-Ebene keine Homologien zueinander auf. So weist Bs3 bspw. Homologie zu Flavin-abhängigen Monooxygenasen (FMOs) auf, während Xa10 vermutlich ein Transmembran-Protein im ER und Xa27 ein kleines Protein

ohne Homologie zu bekannten Resistenzproteine ist (Abb. 4, Römer et al., 2007; Tian et al., 2014; Wu et al., 2008). Letzteres gilt auch für Bs4C-R. Erstaunlicherweise zeigen *in silico* Strukturanalysen, dass Bs4C-R möglicherweise ein TM-Protein ist (Abb. 27) und demnach Xa10 in seiner Funktion ähnlich sein könnte. Auf Sequenzebene besitzen Xa10 und Bs4C-R zwar nur eine Identität von 15 %, jedoch wurden für beide Proteine vier TM-Domänen vorhergesagt (TMHMM-Vorhersageprogramm, Abb. 27) (Tian et al., 2014). Tatsächlich konnten auch YFP-Fusionsproteine von Bs4C-R im ER detektiert werden (Janett Elsaesser, unveröffentlichte Daten).

Für Xa10 wird diskutiert, dass es als Hexamer in der endoplasmatischen Retikulum (ER)-Membran vorliegt und als Calcium-Kanal ( $\text{Ca}^{2+}$ -Kanal) fungiert (Tian et al., 2014). Demnach wäre es auch möglich, dass Bs4C-R einen  $\text{Ca}^{2+}$ -Kanal im ER ausbildet.



**Abb. 27 Proteinvorhersagen für Bs4C-R und Xa10 mit dem Programm TMHMM**

Transmembran-Helices-Vorhersage nach dem Hidden-Markov-Model zeigen, dass Xa10 und Bs4C-R als Multipass-TM-Proteine vorhergesagt sind, wobei die TM-Helices (rote senkrechte Striche), und die Polypeptidkette als Linie in Abhängigkeit von ihrer Position, die entweder innerhalb (blau) oder außerhalb (lila) der Membran liegt, dargestellt.

Xa10 und Bs4C-R induzieren beide Zelltodreaktionen, Xa10 nicht nur in Pflanzenzellen, sondern auch in menschlichen HeLa-Zellen, was vermuten lässt, dass der Mechanismus der Zelltodinduktion konserviert ist (Strauß et al., 2012; Tian et al., 2014). Da Calcium als sekundärer Botenstoff in vielen verschiedenen Zellprozessen wie intrazellulären Signalweiterleitungs-, Genexpression-, und Entwicklungsprozessen eine wichtige Rolle spielt, kann nur spekuliert werden, auf welchem Wege die Zelltodreaktion induziert wird. So führt beispielsweise der Ausstrom von Calciumionen aus dem ER zur Störung der Calcium-Homöostase im Zytoplasma. Dies kann in der Aktivierung von MAPK-Signalkaskaden resultieren, in deren Folge Gene induziert werden, die in Abwehrreaktionen inklusive der Ausbildung der hypersensitiven Reaktion involviert sind (Ma and Berkowitz, 2007).

Es können aber auch wichtige zelluläre Prozesse im ER selbst gestört werden, wenn Calcium-Ionen austreten, wie bspw. die Proteinsekretion, die Lipid-Biosynthese und die Faltung von Proteinen. Ist die Störung der Proteinfaltung dauerhaft, welche auch als *unfolded protein response (UPR)* bezeichnet wird, kann dies ebenfalls eine Zelltodreaktion (Apoptose) verursachen (Shore et al., 2011; Tabas and Ron, 2011). Da für Xa10 lediglich eine Änderung der Calcium-Konzentration untersucht wurde, bleibt offen ob es Kanalaktivität aufweist und als Calciumkanal fungieren kann. Dies müsste durch weitere Analysen, wie bspw. elektrophysiologische Untersuchungen an rekonstituierten Membranen oder Vesikeln, welche Xa10-Oligomere als potentiellen Calciumkanal enthalten, geklärt werden. Jedoch wäre es auch möglich, dass das Membranprotein Xa10 eine andere Funktion aufweist. Für Bs4C-R müsste zuerst mittels Gelfiltration, Ultrazentrifugation oder auch ITC (*Isothermal Titration Calorimetry*) analysiert werden, ob es wie Xa10 Oligomere ausbildet. Die Messung der  $\text{Ca}^{2+}$ -Konzentration im Zytoplasma und ER unter Verwendung von Calcium-Indikatoren kann Hinweise auf eine potentielle Funktion von Bs4C-R als  $\text{Ca}^{2+}$ -Kanal geben, welche durch elektrophysiologische Untersuchungen bestätigt werden müssten.

Interessanterweise kodiert das Allel aus der susceptible *C. pubescens*-Akzession ebenfalls für ein funktionales Protein. So führte die Überexpression von *Bs4C-S*, welches 9 AS- und 2 Indel-Polymorphismen aufweist, verglichen zu *Bs4C-R*, ebenfalls zu einer HR im transienten Assay (siehe 3.5). Und auch die Vorhersage zur Struktur ergab vier putative TM-Domänen und war somit identisch zu der von *Bs4C-R* (siehe A33). Ähnliches ist auch für die *S*-Allele von *Bs3* und *Xa27* bekannt, da *Bs3-E* bei Überexpression eine HR induziert und die KDS von *xa27* identisch zu der von *Xa27* ist (Gu et al., 2005; Römer et al., 2007). Das Vorhandensein der funktionalen Allele in den susceptible Akzessionen könnte auf eine mögliche Funktion außerhalb der Resistenz hinweisen. Interessanterweise spielt Zelltod nicht nur in der Resistenzausbildung sondern auch in Entwicklungsprozessen eine wichtige Rolle (Lam, 2004). Aber für *Bs3* konnte gezeigt werden, dass es nicht in anderen Geweben, wie Wurzel, Blüten und Früchten exprimiert wird und in Blättern nur nach Induktion durch *AvrBs3* (Römer et al., 2009b). Auch *Bs4C-R*- und *Xa10*-Transkript konnte nur nach Induktion durch den entsprechenden TALE detektiert werden. Jedoch ergaben initiale Untersuchungen für *Bs4C-R*, dass es auch in der Wurzel exprimiert wird (Strauß und Lahaye, unveröffentlichte Daten). Da Expressionsstudien über einen

längeren Entwicklungszeitraum der Pflanze noch nicht durchgeführt wurden, bleibt eine mögliche weitere Funktion offen, könnte aber mit Hilfe von transgenen Reporterlinien (bspw. *A. thaliana*) und RT-PCR-Analysen hinterfragt werden.

#### 4.4 Potentielle AvrBs4-Suszeptibilitätsgene in *C. pubescens*

Mit Hilfe der Transkriptomanalysen konnte nicht nur das *R*-Gen *Bs4C-R* sondern auch weitere direkte AvrBs4-induzierte Gene identifiziert werden – darunter Gene, die für ein Methyl-bindendes Protein (MBP) und eine Phosphorylase kodieren (siehe 3.8.1 und 3.8.2). Im Gegensatz zu *Bs4C-R* sind beide Gene sowohl in der resistenten als auch in der suszeptiblen Akzession induziert (siehe 3.8). Ihre vorhergesagten EBEs und kodierende Regionen sind in beiden Akzessionen sequenzidentisch (siehe 3.8). Diese Gene sind im Zusammenhang mit der Virulenz des Bakteriums interessant, da sie potentielle Suszeptibilitätsgene (*S*-Gene) (siehe 1.5) darstellen. Wie erwähnt, sind *S*-Gene Wirtsgene, deren Expression durch TALEs die bakterielle Virulenz begünstigen, indem sie zu einem erhöhten Wachstum oder auch zur Ausbildung von Krankheitssymptomen beitragen (siehe 1.5). Vertreter der *S*-Gene kodieren bspw. für Zucker- oder Sulfattransporter, für eine RNA-Methyltransferase sowie für Transkriptionsfaktoren (Antony et al., 2010; Cernadas et al., 2014; Chen et al., 2010; Hu et al., 2014; Kay et al., 2007; Li et al., 2014; Liu et al., 2011; Moscou and Bogdanove, 2009; Streubel et al., 2013; Sugio et al., 2007; Yang et al., 2006; Yu et al., 2011). Die hier isolierten Kandidaten sind Mitglieder zweier weiterer Proteinfamilien. So kodiert das Gen „630“ für ein Methyl-CpG-bindendes Protein (MBP) für deren Homologe aus *A. thaliana*, welche eine Identität von etwa 50 % zu den *C. pubescens*-MBP besitzen (siehe A9), die Bindung an methylierte (CpG)-Nukleotide *in vitro* nachgewiesen werden (Grafi et al., 2007; Zemach and Grafi, 2003). Es wird angenommen, dass sie als strukturelle Proteine agieren, die Histon-Deacetylasen und Chromatin-Remodeling-Faktoren rekrutieren und somit zur transkriptionellen Repression von Genen beitragen (Ballestar and Wolffe, 2001; Wade, 2001). Im Kontext mit AvrBs4 könnten MBPs Gene reprimieren, die in die Abwehrreaktionen der Pflanze involviert sind und somit dem Bakterium einen Vorteil während der Besiedlung verschaffen.

Der Kandidat 15502 zeigt über 90 % Homologie zu einer alpha-Glucan-Phosphorylase aus Kartoffel, welche in den Kohlenhydrat-Stoffwechsel involviert ist (<http://solgenomics.net/locus/1806/view>) (Brisson et al., 1989).

Alpha-Glucan-Phosphorylasen katalysieren beim Stärkeabbau in Pflanzen die reversible Abspaltung endständiger Glukosemoleküle, indem sie diese phosphorylieren. Das entstandene Glucose-1-phosphat kann in verschiedenen metabolischen Prozessen verarbeitet werden, wobei es teilweise in andere Zellkompartimente transportiert werden muss. Das *SWEET*-Gen *OsSWEET11* kodiert für einen Zuckertransporter, welcher durch den TAL-Effektor PthXo1 aus *Xoo* induziert wird und die Bakterien im Xylem und/oder Apoplasten mit Zucker versorgt. Kann *OsSWEET11* aufgrund von Polymorphismen im EBE nicht mehr induziert werden, führt dies zu einem langsameren Wachstums, welches wahrscheinlich durch die limitierte Zuckerversorgung bedingt ist (Chen et al., 2010). Denkbar wäre, dass die von Kandidat 15502 kodierte Phosphorylase, durch die Bereitstellung von Zuckermolekülen in die Nährstoffversorgung des Bakteriums involviert sein könnte.

Bisherige Analysen zu TAL-induzierten *S*-Genen weisen daraufhin, dass nur ein TAL-Zielgen das eigentliche *S*-Gen ist (Cernadas et al., 2014; Cohn et al., 2014; Hu et al., 2014). Induziert werden aber meist mehr als 90 Zielgene, wie in diversen Transkriptomanalysen nachgewiesen werden konnte (Pereira et al., 2014; Perez-Quintero et al., 2013). Das heißt, es gibt eine große Anzahl an sogenannten *off-targets*, die für das bakterielle Wachstum keine Rolle spielen. Die Identifikation des *S*-Gens unter der Gesamtheit der TAL-Zielgene erfolgt letztendlich über Virulenzanalysen. Für AvrBs4 ist ein positiver Effekt bzgl. der Ausbildung der Krankheitssymptome in Feldexperimenten gezeigt worden (Wichmann und Bergelson 2004). Diese sind im Labor schwer nachzustellen und somit konnte unter Laborbedingungen bislang kein Effekt für AvrBs4 nachgewiesen werden. Der einzige beschriebene Phänotyp für AvrBs4 ist die vermehrte Bildung von Katalasekristalle in Peroxisomen nach transienter Expression in *N. benthamiana* (Gürlebeck et al., 2009). Hauptfunktion der Katalase ist der Abbau des Zellgifts Wasserstoffperoxid, welches unter anderem bei Pathogenbefall vermehrt gebildet wird und somit könnte AvrBs4-induzierte Katalasebildung zur Unterdrückung der Pathogenerkennung beitragen. Ein möglicher Zusammenhang der AvrBs4-induzierten Expression von „630“ und „15502“ bzw. der Funktion beider Genprodukte mit der Bildung dieser

Katalasekristalle ist nicht offensichtlich. Um zu klären, ob eines der beiden Kandidatengene das *S*-Gen darstellt, könnte man dTALEs generieren, welche spezifisch die jeweiligen Kandidaten induzieren und diese in Feldexperimente hinsichtlich der Ausbildung von Krankheitssymptomen analysieren bzw. parallel versuchen, einen Laborassay zu etablieren. Dafür könnten bspw. *Xanthomonas* Typ-III-Effektormutanten wie bspw. *XcvΔavrBs1ΔavrBs2ΔavrBs3ΔavrBs4* oder auch *Xcv* 85-10 *ΔxopBΔxopS* verwendet werden, bei denen eine Reduktion in der Ausbildung der Krankheitssymptome sowie im bakteriellen Wachstum nachgewiesen wurde (Schulze et al., 2012; Wichmann and Bergelson, 2004). Nur der effektordefiziente Stamm mit dem dTALE, der das *S*-Gen induziert, wird einen positiven Effekt auf das bakterielle Wachstum bzw. die Ausbildung von Krankheitssymptomen haben.

#### 4.5 Bs4C-Homologe in *Capsicum*-Funktion außerhalb der Resistenz?

In *C. annuum* cv. CM334 konnten bislang sieben *Bs4C-R*-Homologe *CaBs4C.1-7* identifiziert werden (siehe 3.11.1, A16). Während *CaBs4C.6* und *CaBs4C.7* Homologien von unter 39 % verglichen mit *Bs4C-R* zeigen, weisen *CaBs4C.1-5* eine Identität von mehr als 67 % auf Nukleotidebene gegenüber *Bs4C-R* auf (Tab. 2). Unter diesen fünf Homologen befinden sich jedoch nur drei (*CaBs4C.1*, *CaBs4C.2*, *CaBs4C.5*), die einen durchgängigen Leserahmen besitzen, während in *CaBs4C.3* und *CaBs4C.4* vorzeitige Stoppkodons identifiziert wurden (siehe 3.11.1).

*CaBs4C.1* aus *C. annuum* ist dabei das Ortholog von *Bs4C-R* (89 % Identität) aus *C. pubescens* (siehe 3.11.1). Im funktionalen Test konnte nachgewiesen werden, dass *CaBs4C.1* bei 35S:Überexpression eine HR induziert (siehe 3.11.3). Das Homolog *CaBs4C.2* besitzt eine Insertion von 26 AS (*in frame*) und zusätzliche Polymorphismen verglichen zu *Bs4C-R* (siehe 3.11.1). Ein Vergleich der Strukturvorhersagen für *CaBs4C.2* und *Bs4C-R* zeigt, dass *CaBs4C.2* zwischen der zweiten und dritten TM-Domäne einen größeren Loop aufweist (siehe A34). *CaBs4C.2* induziert bei 35S:Überexpression keine HR, was nahe legt, dass die durch die Duplikation bedingte strukturelle Veränderung, seine Funktionalität beeinflusst (siehe 3.11.3). Interessanterweise konnte in *C. pubescens* *CpBs4C.2-R* identifiziert werden, welches das zu *CaBs4C.2* korrespondierende Ortholog darstellt. Dies ist zu 91 % identisch zu *CaBs4C.2* ohne Duplikation und induziert eine HR bei Überexpression. Für *CpBs4C.2-R* und *CaBs4C.2Del* wird eine ähnliche Struktur mit



drei putativen TM-Domänen vorhergesagt, die vermuten lässt, dass auch CaBs4C.2Del funktional sein könnte (siehe A34).

*Bs4C-R* und *CpBs4C.2-R*, welche zu 85 % identisch sind (A27), können durch Genduplikation in *C. pubescens* entstanden sein, wobei (aufgrund fehlender Evolutions-Untersuchungen) nur spekuliert werden kann, welches Gen das ursprüngliche ist und ob es möglicherweise eine andere Funktion hatte. Nach Genduplikation, ist es möglich, dass 1) beide Gene ihre Funktion behalten und/oder 2) eines in einem anderen Gewebe exprimiert wird, 3) ein funktionsloses Pseudogen entsteht oder 4) das Gen eine andere „Funktion“ übernimmt. Da *CpBs4C.2-R* zwar bei Überexpression eine HR induziert, jedoch nicht durch AvrBs4 induziert wird, könnte es somit ein Pseudogen darstellen. Andererseits, könnte sich *Bs4C-R* von *CpBs4C.2-R* ableiten und durch Polymorphismen in seinem EBE eine neue Funktion in der Erkennung von AvrBs4 und somit in der Resistenz erworben haben. Die unterschiedliche Anzahl von TM-Domänen in Strukturvorhersagen kann auch eine unterschiedliche Lokalisation (Neolokalisation) bzw. Funktion (Neofunktionalisierung) bedingen. So konnte für duplizierte Genpaare aus *A. thaliana* gezeigt werden, dass diese eine unterschiedliche Lokalisationen und einige noch verschiedene Expressionsmuster aufweisen (Liu et al., 2014).

Desweiteren stellt sich die Frage nach der Funktion der Homologen in der Pflanze. Allel-Analysen für *Bs3* zeigten, dass Akzessionen, welche Allele mit vorzeitigen Stoppkodons aufwiesen, in ihren Phänotypen nicht verändert waren. Dies lässt darauf schließen, dass *Bs3* nur eine Funktion im Sinne der Resistenz besitzt. *C. pubescens* und *C. annuum* besitzen mehr als ein *Bs4C*-Allel und viele davon kodieren aufgrund von vorzeitigen Stoppkodons für nicht-funktionale Proteine. Für *Bs4C* konnte jedoch eine Expression in der Wurzel nachgewiesen werden (Tina Strauß und Thomas Lahaye, unveröffentlichte Daten). Somit ließe sich auch für die drei Allele aus *C. annuum* (*CaBs4C.1*, *CaBs4C.2* und *CaBs4C.5*) und das *Bs4C*-Paralog aus *C. pubescens* (*CpBs4C.2-R*), die für Vollständigproteine kodieren, eine Funktion außerhalb der Resistenz vermuten. Für die *Bs4C-R*-Homologen sollten sowohl Expressionanalysen in verschiedenen Geweben (transgene *A. thaliana*) als auch Lokalisationsstudien durchgeführt werden um dies zu hinterfragen.

	Nukleotidposition im <i>EBE</i>																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
<b><i>Bs4C-R</i></b>	T	A	T	A	A	A	A	A	T	A	G	T	C	C	T	C	T	C
<b><i>Bs4C-S</i></b>	T	A	C	C	A	A	A	A	A	T	A	G	T	C	C	T	C	T
<b><i>CaBs4C.3</i></b>	T	G	C	C	A	A	A	A	A	-	T	A	G	C	C	C	T	A
<b><i>CaBs4C.4</i></b>	T	G	C	C	A	A	A	A	A	-	T	A	G	C	C	C	T	C
<b><i>CpBs4C.2-R</i></b>	T	G	C	C	A	A	A	A	A	A	T	A	G	C	C	C	T	C
<b><i>CaBs4C.2</i></b>	T	G	C	C	A	A	A	A	A	A	T	A	G	C	C	C	T	C
<b><i>CaBs4C.1</i></b>	T	A	C	C	A	A	A	A	A	A	T	A	G	C	C	C	T	C
<b><i>CaBs4C.5</i></b>	T	A	C	C	C	G	A	T	A	T	G	T	A	C	A	C	T	A

**Abb 28.: Vergleich der EBE der *Bs4C*-Allele aus *C. pubescens* und *C. annuum***

Alle abgebildeten *EBE* aus *C. annuum* (*CaBs4C.1-5*, grün unterlegt) sowie *Bs4C-S* und *CpBs4C.2-R* aus *C. pubescens* (sowie *Bs4C-R* sind rot unterlegt) zeigen an Position zwei und drei C-Nukleotide (gelb umrandet). Die *EBE* von *CpBs4C.2-R* und *CaBs4C.2* sind identisch. Identische Nukleotide sind weiß markiert und haben einen schwarzen Hintergrund. Unterschiedliche Nukleotide sind schwarz auf weißem Hintergrund dargestellt.

Die Vergleich der putativen *EBE* aus *Capsicum annuum* und *Capsicum pubescens* zeigt, dass nur das *EBE* von *Bs4C-R*, die an Position zwei und drei dem TAL-Code entsprechenden Nukleotide T und A aufweist (Abb. 28). Alle anderen zeigen C-Nukleotide an dieser Position und sind vermutlich nicht durch *AvrBs4* induzierbar. Zusammenfassend lässt sich festhalten, dass die in *C. annuum* identifizierten *Bs4C*-Allele, aufgrund der wahrscheinlich nicht-induzierbaren *EBE* und/oder Polymorphismen in der KDS, nicht-funktionale Allele darstellen. Dies stimmt mit der Beobachtung überein, dass *AvrBs4*-exprimierende Xanthomonaden keine HR auf *C. annuum* cv. CM334 bzw. ECW/ECW-30R induzieren. Interessanterweise konnte auch in anderen *C. annuum*-Akzessionen keine Erkennung von *AvrBs4* beobachtet werden (Patrick Römer und Thomas Lahaye, unveröffentlichte Daten) obwohl Xanthomonaden-Infektionen in *C. annuum* Felder erhebliche Ernteschäden verursachen (Pernezny et al., 1996). Dabei stellt sich natürlich die Frage nach den jeweiligen *Xcv*-Stämme und deren TAL-Effektor Repertoire. Besitzen diese das TAL-Effektorgen *avrBs4* oder vornehmlich *avrBs3*? Diese Frage erfordert weitere Analysen über *Xcv*-Stämme, welche im Feld isoliert wurden. Diese könnten mittels Southern Blot Analysen erfolgen und würde einen Hinweis auf die Anzahl der Effektoren der einzelnen *Xcv*-Stämme geben.

Im Rahmen der Doktorarbeit von Janett Elsaesser werden die Orthologen und Paralogen in unserer Arbeitsgruppe intensiver untersucht.

## 5 Literatur

### LITERATURVERZEICHNIS

Agarwal, M., Singh, A., Mittal, D., Sahi, C., and Grover, A. (2011). Cycloheximide-mediated superinduction of genes involves both native and foreign transcripts in rice (*Oryza sativa* L.). *Plant Physiol Biochem* 49, 9-12.

Agrios, G.N. (1997). *Plant Pathology*, 4th edition edn (San Diego: Academic Press).

Almendral, J.M., Sommer, D., Macdonald-Bravo, H., Burckhardt, J., Perera, J., and Bravo, R. (1988). Complexity of the early genetic response to growth factors in mouse fibroblasts. *Mol Cell Biol* 8, 2140-2148.

Antony, G., Zhou, J., Huang, S., Li, T., Liu, B., White, F., and Yang, B. (2010). Rice *xa13* recessive resistance to bacterial blight is defeated by induction of the disease susceptibility gene *Os-11N3*. *Plant Cell* 22, 3864-3876.

Ballestar, E., and Wolffe, A.P. (2001). Methyl-CpG-binding proteins. Targeting specific gene repression. *European journal of biochemistry / FEBS* 268, 1-6.

Ballvora, A., Schornack, S., Baker, B.J., Ganai, M., Bonas, U., and Lahaye, T. (2001). Chromosome landing at the tomato *Bs4* locus. *Mol Genet Genomics* 266, 639-645.

Boch, J., and Bonas, U. (2010). *Xanthomonas AvrBs3* family-type III effectors: discovery and function. *Annu Rev Phytopathol* 48, 419-436.

Boch, J., Bonas, U., and Lahaye, T. (2014). TAL effectors – pathogen strategies and plant resistance engineering. *New Phytol*, *in press*.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. *Science* 326, 1509-1512.

Bogdanove, A.J. (2014). Principles and applications of TAL effectors for plant physiology and metabolism. *Curr Opin Plant Biol* 19, 99-104.

Bonas, U., Conrads-Strauch, J., and Balbo, I. (1993). Resistance in tomato to *Xanthomonas campestris* pv *vesicatoria* is determined by alleles of the pepper-specific avirulence gene *avrBs3*. *Mol Gen Genet* 238, 261-269.

Bonas, U., Stall, R.E., and Staskawicz, B. (1989). Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol Gen Genet* 218, 127-136.

Booher, N.J., and Bogdanove, A.J. (2014). Tools for TAL effector design and target prediction. *Methods*. *in press*

- Brisson, N., Giroux, H., Zollinger, M., Camirand, A., and Simard, C. (1989). Maturation and subcellular compartmentation of potato starch phosphorylase. *Plant Cell* 1, 559-566.
- Bultmann, S., Morbitzer, R., Schmidt, C.S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T., and Leonhardt, H. (2012). Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res* 40, 5368-5377.
- Burke, D.T., Carle, G.F., and Olson, M.V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236, 806-812.
- Cernadas, R.A., Doyle, E.L., Nino-Liu, D.O., Wilkins, K.E., Bancroft, T., Wang, L., Schmidt, C.L., Caldo, R., Yang, B., White, F.F., *et al.* (2014). Code-assisted discovery of TAL effector targets in bacterial leaf streak of rice reveals contrast with bacterial blight and a novel susceptibility gene. *PLoS Pathog* 10, e1003972.
- Chen, L.-Q., Hou, B.-H., Lalonde, S., Takanaga, H., Hartung, M.L., Qu, X.-Q., Guo, W.-J., Kim, J.-G., Underwood, W., Chaudhuri, B., *et al.* (2010). Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 468, 527-532.
- Chen, L.Q. (2014). SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol* 201, 1150-1155.
- Chen, S., Huang, Z., Zeng, L., Yang, J., Liu, Q., and Zhu, X. (2008). High-resolution mapping and gene prediction of *Xanthomonas Oryzae* pv. *Oryzae* resistance gene Xa7. *Mol Breed* 22, 433-441.
- Cohn, M., Bart, R., Shybut, M., Dahlbeck, D., Gomez, M., Morbitzer, R., Hou, B.H., Frommer, W., Lahaye, T., and Staskawicz, B. (2014). *Xanthomonas axonopodis* virulence is promoted by a transcription activator like (TAL) effector mediated induction of a SWEET sugar transporter in cassava. *Mol Plant Microbe Interact*.
- Cong, L., Zhou, R., Kuo, Y.-c., Cunniff, M., and Zhang, F. (2012). Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nature Communications* 3, 968.
- de Lange, O., Binder, A., and Lahaye, T. (2014a). From dead leaf, to new life: TAL effectors as tools for synthetic biology. *Plant J* 78, 753-771.
- de Lange, O., Schreiber, T., Schandry, N., Radeck, J., Braun, K.H., Koszinowski, J., Heuer, H., Strauss, A., and Lahaye, T. (2013). Breaking the DNA-binding code of *Ralstonia solanacearum* TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. *New Phytol* 199, 773-786.
- de Lange, O., Wolf, C., Dietze, J., Elsaesser, J., Morbitzer, R., and Lahaye, T. (2014b). Programmable DNA-binding proteins from *Burkholderia* provide a fresh perspective on the TALE-like repeat domain. *Nucleic Acids Res* 42, 7436-7449.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y., and Yan, N. (2012). Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* 335, 720-723.

Deslandes, L., Olivier, J., Theulieres, F., Hirsch, J., Feng, D.X., Bittner-Eddy, P., Beynon, J., and Marco, Y. (2002). Resistance to *Ralstonia solanacearum* in *Arabidopsis thaliana* is conferred by the recessive *RRS1-R* gene, a member of a novel family of resistance genes. *Proc Natl Acad Sci USA* 99, 2404-2409.

Dodds, P.N., and Rathjen, J.P. (2010). Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat Rev Genet* 11, 539-548.

Doyle, E.L., Booher, N.J., Standage, D.S., Voytas, D.F., Brendel, V.P., Vandyk, J.K., and Bogdanove, A.J. (2012). TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: tools for TAL effector design and target prediction. *Nucl Acids Res* 40, W117-122.

Franco, A.R., Gee, M.A., and Guilfoyle, T.J. (1990). Induction and superinduction of auxin-responsive mRNAs with auxin and protein synthesis inhibitors. *J Biol Chem* 265, 15845-15849.

Gassmann, W., Hinsch, M.E., and Staskawicz, B.J. (1999). The *Arabidopsis* *RPS4* bacterial-resistance gene is a member of the TIR-NBS-LRR family of disease-resistance genes. *Plant J* 20, 265-277.

Grafi, G., Zemach, A., and Pitto, L. (2007). Methyl-CpG-binding domain (MBD) proteins in plants. *Biochim Biophys Acta* 1769, 287-294.

Grau, J., Boch, J., and Posch, S. (2013). TALENoffer: genome-wide TALEN off-target prediction. *Bioinformatics* 29, 2931-2932.

Gu, K., Sangha, J.S., Li, Y., and Yin, Z. (2008). High-resolution genetic mapping of bacterial blight resistance gene *Xa10*. *Theor Appl Genet* 116, 155-163.

Gu, K., Yang, B., Tian, D., Wu, L., Wang, D., Sreekala, C., Yang, F., Chu, Z., Wang, G.L., White, F.F., *et al.* (2005). R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* 435, 1122-1125.

Gürlebeck, D. (2001). Genetische und molekulare Analyse von AvrBs3 und AvrBs4, zwei Mitgliedern der AvrBs3-Genfamilie aus *Xanthomonas campestris* pv. *vesicatoria*. In Institut für Genetik (Halle, Germany: Martin-Luther Universität Halle-Wittenberg).

Gürlebeck, D., Raschke, A., Kirchner, O., Hause, G., and Bonas, U. (2009). Visualization of novel virulence activities of the *Xanthomonas* type III effectors AvrBs1, AvrBs3 and AvrBs4. *Mol Plant-Microbe Interact* 10, 175-188.

He, P., Shan, L., and Sheen, J. (2007). Elicitation and suppression of microbe-associated molecular pattern-triggered immunity in plant-microbe interactions. *Cell Microbiol* 9, 1385-1396.

Horvath, D.M., Huang, D.J., and Chua, N.H. (1998). Four classes of salicylate-induced tobacco genes. *Mol Plant Microbe Interact* 11, 895-905.

- Hu, Y., Zhang, J., Jia, H., Sosso, D., Li, T., Frommer, W.B., Yang, B., White, F.F., Wang, N., and Jones, J.B. (2014). Lateral organ boundaries 1 is a disease susceptibility gene for citrus bacterial canker disease. *Proc Natl Acad Sci USA* *111*, E521-529.
- Hummel, A.W., Doyle, E.L., and Bogdanove, A.J. (2012). Addition of transcription activator-like effector binding sites to a pathogen strain-specific rice bacterial blight resistance gene makes it effective against additional strains and against bacterial leaf streak. *New Phytol* *195*, 883-893.
- Iyer, A.S., and McCouch, S.R. (2004). The rice bacterial blight resistance gene xa5 encodes a novel form of disease resistance. *Mol Plant Microbe Interact* *17*, 1348-1354.
- Jankele, R., and Svoboda, P. (2014). TAL effectors: tools for DNA Targeting. *Briefings in functional genomics* *13*, 409-419.
- Jones, J.B., Stall, R.E., and Bouzar, H. (1998). Diversity among xanthomonads pathogenic on pepper and tomato. *Annu Rev Phytopathol* *36*, 41-58.
- Jones, J.D., and Dangl, J.L. (2006). The plant immune system. *Nature* *444*, 323-329.
- Jordan, T., Romer, P., Meyer, A., Szczesny, R., Pierre, M., Piffanelli, P., Bendahmane, A., Bonas, U., and Lahaye, T. (2006). Physical delimitation of the pepper Bs3 resistance gene specifying recognition of the AvrBs3 protein from *Xanthomonas campestris* pv. *vesicatoria*. *Theor Appl Genet* *113*, 895-905.
- Jordan, T. (2005). Genetische und physikalische Limitierung des Bs3 Resistenzgen-Locus in *Capsicum annuum*. In Institut für Genetik (Halle, Germany: Martin-Luther Universität Halle-Wittenberg).
- Kay, S., Hahn, S., Marois, E., Hause, G., and Bonas, U. (2007). A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science* *318*, 648-651.
- Kay, S., Hahn, S., Marois, E., Wieduwild, R., and Bonas, U. (2009). Detailed analysis of the DNA recognition motifs of the *Xanthomonas* type III effectors AvrBs3 and AvrBs3Deltarep16. *Plant J* *59*, 859-871.
- Kim, S., Park, M., Yeom, S.I., Kim, Y.M., Lee, J.M., Lee, H.A., Seo, E., Choi, J., Cheong, K., Kim, K.T., *et al.* (2014). Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. *Nat Genet* *46*, 270-278.
- Lam, E. (2004). Controlled cell death, plant survival and development. *Nat Rev Mol Cell Biol* *5*, 305-315.
- Leoni, C., Volpicella, M., De Leo, F., Gallerani, R., and Ceci, L.R. (2011). Genome walking in eukaryotes. *The FEBS journal* *278*, 3953-3977.
- Li, Z., Zou, L., Ye, G., Xiong, L., Ji, Z., Zakria, M., Hong, N., Wang, G., and Chen, G. (2014). A potential disease susceptibility gene CsLOB of citrus is targeted by a

major virulence effector PthA of *Xanthomonas citri* subsp. *citri*. *Mol Plant* 7, 912-915.

Liu, Q., Yuan, M., Zhou, Y.A.N., Li, X., Xiao, J., and Wang, S. (2011). A paralog of the MtN3/saliva family recessively confers race-specific resistance to *Xanthomonas oryzae* in rice. *Plant Cell Environ* 34, 1958-1969.

Ma, W., and Berkowitz, G.A. (2007). The grateful dead: calcium and cell death in plant innate immunity. *Cell Microbiol* 9, 2571-2585.

Mak, A.N., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* 335, 716-719.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S.V., Machado, M.A., *et al.* (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol* 13, 614-629.

Marois, E., Van den Ackerveken, G., and Bonas, U. (2002). The xanthomonas type III effector protein AvrBs3 modulates plant gene expression and induces cell hypertrophy in the susceptible host. *Mol Plant Microbe Interact* 15, 637-646.

McHale, L., Tan, X., Koehl, P., and Michelmore, R.W. (2006). Plant NBS-LRR proteins: adaptable guards. *Genome Biol* 7, 212.

Meckler, J.F., Bhakta, M.S., Kim, M.S., Ovadia, R., Habrian, C.H., Zykovich, A., Yu, A., Lockwood, S.H., Morbitzer, R., Elsaesser, J., *et al.* (2013). Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res* 41, 4118-4128.

Minsavage, G., Jones, J., Stall, R., Miller, S., and Ritchie, D. (1999). Hypersensitive resistance in *Capsicum pubescens* PI 235047 to *Xanthomonas campestris* pv. *vesicatoria* (Xcv) is elicited by *avrBs3-2*. *Phytopathology* 89, S53.

Morbitzer, R., Elsaesser, J., Hausner, J., and Lahaye, T. (2011). Assembly of custom TALE-type DNA binding domains by modular cloning. *Nucl Acids Res* 39, 5790-5799.

Morbitzer, R., Römer, P., Boch, J., and Lahaye, T. (2010). Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc Natl Acad Sci USA* 107, 21617-21622.

Moscou, M.J., and Bogdanove, A.J. (2009). A simple cipher governs DNA recognition by TAL effectors. *Science* 326, 1501.

Nino-Liu, D.O., Ronald, P.C., and Bogdanove, A.J. (2006). *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Mol Plant Pathol* 7, 303-324.

Noel, L.D., Denance, N., and Szurek, B. (2013). Predicting promoters targeted by TAL effectors in plant genomes: from dream to reality. *Front Plant Sci* 4, 333.

Pereira, A.L., Carazzolle, M.F., Abe, V.Y., de Oliveira, M.L., Domingues, M.N., Silva, J.C., Cernadas, R.A., and Benedetti, C.E. (2014). Identification of putative TAL effector targets of the citrus canker pathogens shows functional convergence underlying disease development and defense response. *BMC Genomics* 15, 157.

Perez-Quintero, A.L., Rodriguez, R.L., Dereeper, A., Lopez, C., Koebnik, R., Szurek, B., and Cunnac, S. (2013). An improved method for TAL effectors DNA-binding sites prediction reveals functional convergence in TAL repertoires of *Xanthomonas oryzae* strains. *PLoS One* 8, e68464.

Pernezny K., Datnoff, L., Mueller, T. and Collins, J. (1996). Effect of bacterial spot on yield and quality of fresh market tomatoes. *HortScience* 18:69–70

Peters, J.L., Cnudde, F., and Gerats, T. (2003). Forward genetics and map-based cloning approaches. *Trends Plant Sci* 8, 484-491.

Pierre, M., Noël, L., Lahaye, T., Ballvora, A., Veuskens, J., Ganai, M., and Bonas, U. (2000). High-resolution genetic mapping of the pepper resistance locus *Bs3* governing recognition of the *Xanthomonas campestris* pv. *vesicatora* AvrBs3 protein. *Theor Appl Genet* 101, 255-263.

Plessy, C., Bertin, N., Takahashi, H., Simone, R., Salimullah, M., Lassmann, T., Vitezic, M., Severin, J., Olivarius, S., Lazarevic, D., *et al.* (2010). Linking promoters to functional transcripts in small samples with nanoCAGE and CAGEscan. *Nat Methods* 7, 528-534.

Römer, P. (2010). Isolierung des Paprika Bs3-Resistenzgens und Interaktionsanalyse zwischen TAL-Effektoren und pflanzlichen Promotoren. In Institut für Genetik (Halle, Germany: Martin-Luther Universität Halle-Wittenberg).

Römer, P., Hahn, S., Jordan, T., Strauß, T., Bonas, U., and Lahaye, T. (2007). Plant-pathogen recognition mediated by promoter activation of the pepper *Bs3* resistance gene. *Science* 318, 645-648.

Römer, P., Recht, S., and Lahaye, T. (2009a). A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens. *Proc Natl Acad Sci USA* 106, 20526-20531.

Römer, P., Recht, S., Strauß, T., Elsaesser, J., Schornack, S., Boch, J., Wang, S., and Lahaye, T. (2010a). Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytol* 187, 1048-1057.

Römer, P., Recht, S., Strauß, T., Elsaesser, J., Schornack, S., Boch, J., Wang, S., and Lahaye, T. (2010b). Promoter elements of rice susceptibility genes are bound and activated by specific TAL effectors from the bacterial blight pathogen, *Xanthomonas oryzae* pv. *oryzae*. *New Phytol* 187, 1048-1057.

Römer, P., Strauß, T., Hahn, S., Scholze, H., Morbitzer, R., Grau, J., Bonas, U., and Lahaye, T. (2009b). Recognition of AvrBs3-like proteins is mediated by specific binding to promoters of matching pepper *Bs3* alleles. *Plant Physiol* 150, 1697-1712.



Ruan, Y.L. (2014). Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu Rev Plant Biol* 65, 33-67.

Salimullah, M., Sakai, M., Plessy, C., and Carninci, P. (2011). NanoCAGE: a high-resolution technique to discover and interrogate cell transcriptomes. *Cold Spring Harb Protoc* 2011, pdb prot5559.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A laboratory Manual*, 2nd edn (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Schaller, A. (2002) Die Abwehr von Frassfeinden: Selbstverteidigung im Pflanzenreich. *Vierteljahresschrift der Naturforschenden Gesellschaft in Zürich*, 147/4, 141-150.

Schornack, S. (2006). Struktur, Erkennungsspezifität und Regulationsmechanismus des Resistenzproteins Bs4 aus Tomate. In Institut für Genetik (Halle, Germany: Martin-Luther Universität Halle-Wittenberg).

Schornack, S., Ballvora, A., Gurlebeck, D., Peart, J., Baulcombe, D., Ganai, M., Baker, B., Bonas, U., and Lahaye, T. (2004). The tomato resistance protein Bs4 is a predicted non-nuclear TIR-NB-LRR protein that mediates defense responses to severely truncated derivatives of AvrBs4 and overexpressed AvrBs3. *Plant J* 37, 46-60.

Schornack, S., Moscou, M.J., Ward, E.R., and Horvath, D.M. (2013). Engineering plant disease resistance based on TAL effectors. *Annu Rev Phytopathol* 51, 383-406.

Schornack, S., Peter, K., Bonas, U., and Lahaye, T. (2005). Expression levels of avrBs3-like genes affect recognition specificity in tomato Bs4- but not in pepper Bs3-mediated perception. *Mol Plant Microbe Interact* 18, 1215-1225.

Schulze, S., Kay, S., Buttner, D., Egler, M., Eschen-Lippold, L., Hause, G., Kruger, A., Lee, J., Muller, O., Scheel, D., *et al.* (2012). Analysis of new type III effectors from *Xanthomonas* uncovers XopB and XopS as suppressors of plant immunity. *New Phytol* 195, 894-911.

Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci USA* 89, 8794-8797.

Shore, G.C., Papa, F.R., and Oakes, S.A. (2011). Signaling cell death from the endoplasmic reticulum stress response. *Curr Opin Cell Biol* 23, 143-149.

Stall, R.E., Jones, J.B., and Minsavage, G.V. (2009). Durability of resistance in tomato and pepper to xanthomonads causing bacterial spot. *Annu Rev Phytopathol* 47, 265-284.

Strauß, T., van Poecke, R.M., Strauß, A., Romer, P., Minsavage, G.V., Singh, S., Wolf, C., Strauß, A., Kim, S., Lee, H.A., *et al.* (2012). RNA-seq pinpoints a *Xanthomonas* TAL-effector activated resistance gene in a large-crop genome. *Proc Natl Acad Sci USA* 109, 19480-19485.

Streubel, J., Pesce, C., Hutin, M., Koebnik, R., Boch, J., and Szurek, B. (2013). Five phylogenetically close rice SWEET genes confer TAL effector-mediated susceptibility to *Xanthomonas oryzae* pv. *oryzae*. *New Phytol* 200, 808-819.

Sugio, A., Yang, B., Zhu, T., and White, F.F. (2007). Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes OsTFIIAgamma1 and OsTFX1 during bacterial blight of rice. *Proc Natl Acad Sci USA* 104, 10720-10725.

t Hoen, P.A.C., Ariyurek, Y., Thygesen, H.H., Vreugdenhil, E., Vossen, R.H.A.M., de Menezes, R.X., Boer, J.M., van Ommen, G.J.B., and den Dunnen, J.T. (2008). Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucl Acids Res* 36.

Tabas, I., and Ron, D. (2011). Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nat Cell Biol* 13, 184-190.

Tian, D., Wang, J., Zeng, X., Gu, K., Qiu, C., Yang, X., Zhou, Z., Goh, M., Luo, Y., Murata-Hori, M., *et al.* (2014). The rice TAL effector-dependent resistance protein XA10 triggers cell death and calcium depletion in the endoplasmic reticulum. *Plant Cell* 26, 497-515.

Upadhyaya, N.M., Mago, R., Staskawicz, B.J., Ayliffe, M.A., Ellis, J.G., and Dodds, P.N. (2014). A bacterial type III secretion assay for delivery of fungal effector proteins into wheat. *Mol Plant Microbe Interact* 27, 255-264.

Van den Ackerveken, G., Marois, E., and Bonas, U. (1996). Recognition of the bacterial avirulence protein AvrBs3 occurs inside the host plant cell. *Cell* 87, 1307-1316.

Van Larebeke, N., Engler, G., Holsters, M., Van den Elsacker, S., Zaenen, I., Schilperoort, R.A., and Schell, J. (1974). Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature* 252, 169-170.

Wade, P.A. (2001). Methyl CpG-binding proteins and transcriptional repression. *BioEssays* 23, 1131-1137.

Wichmann, G., and Bergelson, J. (2004). Effector genes of *Xanthomonas axonopodis* pv. *vesicatoria* promote transmission and enhance other fitness traits in the field. *Genetics* 166, 693-706.

Wu, L., Goh, M.L., Sreekala, C., and Yin, Z. (2008). XA27 depends on an amino-terminal signal-anchor-like sequence to localize to the apoplast for resistance to *Xanthomonas oryzae* pv *oryzae*. *Plant Physiol* 148, 1497-1509.

Xiang, T., Zong, N., Zou, Y., Wu, Y., Zhang, J., Xing, W., Li, Y., Tang, X., Zhu, L., Chai, J., *et al.* (2008). *Pseudomonas syringae* effector AvrPto blocks innate immunity by targeting receptor kinases. *Curr Biol* 18, 74-80.

Yang, B., Sugio, A., and White, F.F. (2006). *Os8N3* is a host disease-susceptibility gene for bacterial blight of rice. *Proc Natl Acad Sci USA* 103, 10503-10508.

Yu, Y., Streubel, J., Balzergue, S., Champion, A., Boch, J., Koebnik, R., Feng, J., Verdier, V., and Szurek, B. (2011). Colonization of rice leaf blades by an African strain of *Xanthomonas oryzae* pv. *oryzae* depends on a new TAL effector that induces the rice nodulin-3 Os11N3 gene. *Mol Plant Microbe Interact* 24, 1102-1113.

Zemach, A., and Grafi, G. (2003). Characterization of *Arabidopsis thaliana* methyl-CpG-binding domain (MBD) proteins. *Plant J* 34, 565-572.

Zhou, J.-M., and Chai, J. (2008). Plant pathogenic bacterial type III effectors subdue host responses. *Curr Opin Microbiol* 11, 179-185.

Zipfel, P.F., Irving, S.G., Kelly, K., and Siebenlist, U. (1989). Complexity of the primary genetic response to mitogenic activation of human T cells. *Mol Cell Biol* 9, 1041-1048.

Zur, H., and Tuller, T. (2013). New universal rules of eukaryotic translation initiation fidelity. *PLoS Comput Biol* 9, e1003136.

## 6 Anhang

### A 1

Abbildung modifiziert nach Strauß et al., 2012 Tab.S2.

Tab. 1 Bst4C-Kandidaten Cluster nach Illumina tag-profiling<sup>1</sup>

Pflanzen-Genotyp Inokulation mit	<i>C. pubescens</i> PI 235047										<i>C. pubescens</i> PI 585270										Max C <sup>2</sup>	PI 235047 X <sub>Cv</sub> <sup>AvBst4</sup> 24 hpi / Max X <sub>Cv</sub>																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
	C	6	12	18	24	X <sub>Cv</sub>	C	6	12	18	24	X <sub>Cv</sub>	C	6	12	18	24	X <sub>Cv</sub> <sup>AvBst4</sup>																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Cluster <sup>3</sup>	hpi																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		

<sup>1</sup> Zahlen entsprechen der Anzahl der Sequenzen pro Cluster und Probe normalisiert gegen die Unterschiede der gesamten gefilterten Anzahl der Sequenzen pro Probe.

<sup>2</sup> Der Wert Max C in der vorletzten Spalte gibt die maximale Anzahl von Reads pro Cluster in den Kontrollproben (*C. pubescens* PI 235047/PI 585270 inokuliert mit X<sub>Cv</sub>) an.

<sup>3</sup> Die Cluster wurden nach der letzten Tabellenspalte geordnet, die das Verhältnis der maximalen reads in den X<sub>Cv</sub><sup>AvBst4</sup> PI 235047 24 hpi-Proben und den der Kontrollproben (siehe<sup>2</sup>) angibt.

A 2

Tab. 2 Bs4C Kandidaten Cluster nach Illumina tag-profiling (Tabellen 1-3)

Pflanzen Genotyp		C. pubescens PI 235047										C. pubescens PI 585270									
Infiltration mit	K	6	12	18	24	Xcv	6	12	18	24	Xcv	K	6	12	18	24	Xcv	6	12	18	24
hpi	0	6	12	18	24	Xcv	6	12	18	24	Xcv	0	6	12	18	24	Xcv	6	12	18	24
<b>Teil 1</b>																					
Candidate1	0	3	3	10	4		3	8	35	611		0	0	0	0	0	0	0	0	0	0
Candidate2	0	0	11	2	1		3	5	6	157		0	0	0	0	0	0	0	0	0	0
Candidate3	0	2	4	2	1		4	5	1	128		0	0	0	0	0	0	0	0	0	0
Candidate4	3	2	1	12	3		4	2	14	397		2	0	1	3	4	0	0	0	0	15
Candidate5	0	0	0	2	0		1	4	3	301		0	0	0	1	4	10	0	1	3	27
Candidate6	1	0	0	1	0		4	0	1	322		0	0	0	0	12	17	0	0	3	28
Candidate7	0	0	0	0	0		0	1	0	110		0	0	0	0	6	0	0	0	0	11
Candidate8	0	0	0	0	0		0	0	0	106		0	0	0	0	8	0	0	0	0	1
Candidate9	8	16	2	3	1		5	3	1	125		3	6	7	6	10	10	3	3	3	1
Candidate10	0	1	5	1	2		1	1	8	211		0	0	1	3	19	0	0	0	1	17
Candidate11	0	2	3	1	1		1	3	5	156		0	1	1	2	15	0	1	1	1	24
Candidate12	0	2	0	3	1		1	2	4	135		0	0	0	11	29	0	1	3	23	95
Candidate13	0	0	1	4	2		0	4	3	356		0	0	0	61	82	0	3	23	95	6
Candidate14	0	1	0	0	1		1	1	1	113		0	0	1	19	53	1	0	2	51	4
Candidate15	0	1	1	2	1		1	3	3	198		0	1	3	66	109	0	5	21	91	1
Candidate16	0	1	0	0	0		0	0	2	141		0	1	3	22	123	0	5	15	158	0
<b>Teil 2</b>																					
Cluster15502	1	1	2	0	1		0	0	14	11		0	2	0	3	0	0	1	2	22	18
Cluster11324	0	0	0	2	3		2	2	5	91		0	0	0	0	0	0	0	0	0	0
Cluster13224	0	0	0	0	0		0	1	0	77		0	0	0	0	4	0	0	0	0	7
Cluster3588	0	1	0	4	4		1	2	8	327		0	0	0	0	0	0	0	0	0	0
Cluster8248	0	0	0	1	0		0	2	14	27		0	1	0	5	0	0	0	9	44	53
Cluster14681	1	4	3	3	1		4	4	5	15		1	5	3	4	1	1	4	2	6	6
Cluster15146	1	4	1	1	1		1	3	5	16		1	5	4	5	4	4	3	5	6	4
Cluster14486	0	0	0	0	0		0	0	0	74		0	0	0	0	5	0	0	0	0	1
Cluster10771	3	2	1	4	1		2	2	25	16		1	0	0	5	5	0	1	25	19	0
<b>Teil 3</b>																					
Contig1545	0	0	2	5	8		4	6	8	437		0	0	0	0	0	0	0	0	0	0
Cluster3000	0	0	6	8	4		2	13	51	323		0	0	0	0	0	0	0	0	0	0
Contig1716	0	2	11	12	10		5	23	25	239		0	0	0	0	0	0	0	0	0	0
Cluster1923	0	10	19	11	23		3	51	18	456		0	0	0	0	0	0	0	0	0	0
Contig1919	3	6	4	1	6		5	3	2	131		3	5	5	4	9	7	3	9	8	17
Cluster8984	0	1	4	5	2		2	1	4	86		0	2	0	7	7	0	3	2	0	6
Cluster13571	0	6	0	1	1		3	3	1	67		0	0	0	0	0	0	0	0	0	0
Cluster699	0	2	6	84	30		1	13	137	1006		0	0	4	43	60	1	3	29	49	0
Cluster11620	0	0	0	2	1		0	3	3	60		2	4	2	5	6	3	2	4	4	4
Cluster9054	0	0	0	0	0		0	0	1	1		0	3	4	3	4	3	10	51	59	0
Contig630	6	13	2	4	7		4	7	185	132		4	21	8	46	7	12	14	349	244	0

Zahlen entsprechen der normalisierten Anzahl der Sequenzen pro Cluster und Probe  
K - Kontrolle, nicht infiltriertes Material der jeweiligen Akzession  
hpi (hours post inoculation)-Stunden nach Infiltration

## A 3

Abbildung modifiziert nach Strauß et al., 2012 Tab.S1.

<b>Tab. 3 Analyse der <i>C. pubescens</i> F<sub>2</sub>-Segreganten, die für die genetische Kartierung des <i>Bs4C</i>-Kandidatengens 12600 verwendet wurden.</b>			
Pflanze	<i>Xcv</i> <sup>avrBs4</sup> Phänotyp	<i>Bs4C</i> -CAPS Marker	Bezeichnung in Abbildung Ergebnisteil
PI 235047	HR	R	R
PI 585270	keine HR	S	S
171-1	HR	R	1
171-2	keine HR	S	2
171-3	HR	R	3
171-4	HR	R	5
171-5	keine HR	S	4
171-6	keine HR	S	6
171-7	HR	R*	
171-8	keine HR	S	8
171-9	HR	R	7
171-10	keine HR	S	10
171-11	HR	R	9
171-12	HR	R*	17
173-1	keine HR	S	12
173-2	HR	R	11
173-3	HR	R*	15
173-4	HR	R*	16
173-5	HR	R	13
173-6	HR	R	
173-7	HR	R	
173-8	keine HR	S	14
173-9	HR	R	
173-10	HR	R*	
173-11	keine HR	S	
173-12	HR	R	
HR - hypersensitive Reaktion S - susceptibler Genotyp R - resistenter Genotyp (homozygot) R* - resistenter Genotyp (heterozygot) PI 235047 – AvrBs4-resistenter <i>C. pubescens</i> -Elter PI 585270 – AvrBs4-susceptibler <i>C. pubescens</i> -Elter			

Putative Start- und Stoppkodon sind mit dick-gedruckten grünen und roten Buchstaben markiert. Grau-hinterlegte Buchstaben zeigen die putative kodierende Sequenz der Gene. Der mittels RACE-Analysen ermittelte Transkriptionsstart für *Bs4C-R* ist durch einen fett gedruckten blauen Buchstaben dargestellt. Das funktionale Effektorbindeelement *EBE<sub>AvrBs4</sub>*-*Bs4C-R* des *Bs4C-R* Promoter ist gelb hinterlegt, die korrespondierende Sequenz

>Bs4C-R (*Capsicum pubescens*; PI 235047; AvrBs4-induzierbar)

TAAATTAACAGTGTTAATTAATAAGAAAGCAACTCTATCCACTGCCGCCGAGTATCTTTTAGGGATTGTTTGTGTGAGGTATAA  
GTAATCTCTGGGATAAAATATAAAGATCAACATTTATCCCATGTTTGTATCAGAGGGATTAACTAGTGGTGAGATAACTTATCCCCCAT  
TACACATAGTGTGATGAATAACTTTATCCCATATTCACAGTGAGATAAGTTATCCCAAAAAATCCCAATATTAAATATCCAAAAAT  
ATAATTTCCCAACCAACGAACCTTTTAAAGTTAAGAGACTATAGGAAACGCAAGATTGACTTATGGAGAGTTCTAAGTCCGAAA  
TTGTTTGTGTCAAATGGGGCCTACGAAAAACATATTATTCTTCGCAAGAACAACTAGCTTACACTACACTAACGTCGCCAGTGGAAAA  
GTACCTTCTATGCAGCAGTATGGTACAATATTATAAAGCGCTGGAACATTTTAAATataaataAGATGCTCTCTCAGGCCCTAGACAA  
GCAGGAATAAAATTTCTAAATCAACTTTCGGATTGGTATCTTCGAGGAAGAGCACCATTAAATAGCAGTGCTGTATTAAT  
AAATACACAACCAGCATACTACACCAACCATAGAAATTTTAGGAAAAGATTCAAAGAATAGAATGAGAGTTTGATCTCAGATACTT  
GATCTTGATTCTTGGCTAACATAGCTCCTCAATCAATATTATCCATTTCTGATAACTGGGATCCTTCCCATATATTTCATGACCATCCCA  
GTTTCATCGTCTTCTCAATCAAAAGCTCTTCTTCTTTTCATATTTCCTTTATTTCTCCATCACTCGTATAACACCTCATCATCCCA  
AATATACGAATACGTGTACGTACTACTACTCTCAGCAGACTCTCCAAAGTCCTTTAATATCTGTGTCTAGCTTCTCTCTACTCCC  
ACAAATGTTGTTCTGGTACTTTTTCGTCATTATCATTTGCACCTTCCCTCATGTTCTTCTTGATTGGCAATACATGGGCTAGTTTTC  
GACAACGGATTTGTCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATGATGGAGTTTGACAGAAAATAATGTG  
AGCCAGCAACATGAAACTCAACAAATCGCTGCTCTAGTGTGAATTCATCGATCAAGGATAATTTTATGTTTCATGTTCCCAAGCTC  
TTGATTTTCATATATGTGTGTTTTTCTGTATGAATTAGAAATTAAGGTACGTTTTCGCTACTATAACAGGTAAAAGATGTCAT  
GACAGTGTGAAATTTATGGGTGCTATTACATGGACTACTCTTTTATACCAATATAGATATAGTTGTTTAGCGATGTGTATACATA  
AATGGATACATGAGCTTTGTCTGCTTGTATGGATATTATCATCTGCGAAGCATTTTCCATGTTAACTATTGTGTTGAATGCTGCT  
GATACGTGTTAATTAGTCACTTGTTAGTTTGGAGTAAAGCTTTTCTATAATATTCTGTGTTATTACAAGTATTTCAAGTTTGTAC  
TAATTTCAATTTCTCTCTCTCTCTCTCTCTG

>Bs4C-S (*Capsicum pubescens*; PI 585270; nicht AvrBs4-induzierbar)

TAAATTAAC TAGTTAATTA AAAAAAGCAACTCTATCCACTGCCACCGCAGTATCTTTTAGGGATTGTTTAGTGTGAGGTATAA  
GTAATCTCGGATAAAATATAAAGATCAATTATCCCATGTTTGATTGAAGGGATTAATCAATGGTGAGATAACTTATCCCTATT  
TACACCATAGTGATAGAATAACTATTCCATATTCATAGTAGAGATAAGTTATCCCAAAAAATCTCAATATTTTATTCAAAAATA  
ATAATTTCCCAACCAACGAACTATGTTAAGATTTCCACAGATGACTTATGGAGAGCTCTAAGTCCGACATGTTTTCGTCAAAAT  
GGGGCCTACGAAAAACATTATTTCTTCGCAGGACACAAC TAGTTTGACTACACTAACTTCCAGTGGAAGATACCTTCATATAGC  
ATTATGGTACAAATATTATTAAGCGCAGGAACATTTTAA **TACcAAAAATAGTCTCTC** AGCCCTCAGACAAGTAGGATAAAAAT  
ATTCTAAATCAACTTTGGATTGGATCTCGAGGAAGAGCCACCTATTTAAATGTATAATAACACACACCAGCATACAC  
CCAACCATAGAAATTTTAGGAAAAGATTCAACAGAAATAAT **ATG**AGGTGGATCTCAGATCTCATCTTAAATTTGGCTAACATG  
ATCAATCAATATTATCATCCATTTCTGATAACTGGGATCCTTTCCATATATTTTCATGACCATCCAGTTTCATCGTCTTCATCAA  
TAAGCTCTTCTCTTTTCATATTTCTCTTTTCTTTCTCCATCATCGTATTAACATTTTCATCATCCAAATATGCAATACGTTATAC  
GTACTACTCTTTCAGAGATCTCTCAAGTCCTTAAATATCTGTGCTAGCTTCTCTTACTCTCCACAAATGTTGTTCTGGTAT  
TTTTTCTTCATCACCATTACACTTTCTCATGTTCTTCTTGATTGGCAATGCATGGGCTAGTTTTCGAGGAAGGACTCTGCATAT  
TTTCTCTACAATATTTCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAAATATGTGTAGGCAGCAACATGAAACTCAA  
CAAAATCGCCTGCTCGTCT **TAG**TTGAAATGAATTCATGCATCAAGGATATATTTTCATGTTCTCAACAGCTCTGTATTTCATATATGT  
GTGTTGTTCTCTGATTTGAATTAATAAAGTAAGTAGCTTTAGCTACTATAACAGGTAAGAGTGTATCAGAGTGTGAAATTTAT  
GGGTGTCTATTACATGGACTACTCTTTTATACCAATAT **AGATCT**ATCTGTTTAGCGATGTGTATACATAAATGGATACATGAGCTT  
TGTCACTTGTATGGATATTCATCATCTGCCGAAGCATTTTCATGTTAACTATTGTGTTGAATGCTGCTGATCTGTTAATTAGT  
CACTTGTATGTTTGTAGTAAAGTTTCTTATAATATTTCTGTTGTTATTATAAGTATTTCTAAGTTTGTACTAATTTCAATTTCTCTTC  
TCTTCTCTCTTG

Vergleich der Nukleotidsequenzen der *C. pubescens* Gene *Bs4C-R* (PI 235047; AvrBs4-induzierbar) und *Bs4C-S* (PI 585270; nicht AvrBs4-induzierbar).  
Abbildung modifiziert nach Strauß et al., 2012 SI Text 3 und SI Text 5.

Das Alignment wurde mit den Programmen ClustalW und Boxshade erstellt. Identische Nukleotide sind als weiße Buchstaben auf schwarzem Hintergrund dargestellt. Ein Strich (-) markiert eine Lücke im Alignment. Putative Start- und Stoppkodon sind mit dick-gedruckten grünen und roten Buchstaben markiert. Grau-hinterlegte Buchstaben zeigen die putative kodierende Sequenz der Gene. Der mittels RACE-Analysen ermittelte Transkriptionsstart für *Bs4C-R* ist durch einen fett gedruckten blauen Buchstaben dargestellt. Das funktionale Effektorbindeelement *EBE<sub>AvrBs4</sub>Bs4C-R* des *Bs4C-R* Promoter ist gelb hinterlegt, die korrespondierende Sequenz *EBE<sub>AvrBs4</sub>Bs4C-S* des *Bs4C-S* Promoters ebenso. Polymorphismen in beiden Effektorbindeelementen sind durch kleine fett-gedruckte Buchstaben darstellt. Die *Bgl*II-Restriktionsschnittstelle (AGATCT) in *Bs4C-S*, welche zur Kartierung auf den F<sub>2</sub>-Segreganten genutzt wurde, ist blau hinterlegt. Die korrespondierende Sequenz in *Bs4C-R* ebenso, wobei ein Polymorphismus in der Sequenz, der mit einem



kleinen Buchstaben markiert ist, die Restriktion verhindert. Die kursiv dargestellte Sequenz kennzeichnet die Region, die für das genetische Mapping amplifiziert wurde.

<i>Bs4C-R</i>	1	TAAATTAACTAGTTAATTAAAAAGAAGCAACTCTATCCACTGCCGCCGCAGTATCTTT
<i>Bs4C-S</i>	1	TAAATTAACTAGTTAATTAAAAAGAAGCAACTCTATCCACTGCCGCCGCAGTATCTTT
<i>Bs4C-R</i>	61	TAGGGATTGTTTGTGAGGTATAAGTAATTCTGGGATAAAATATAAAGATCAATTTAT
<i>Bs4C-S</i>	61	TAGGGATTGTTTGTGAGGTATAAGTAATTCTGGGATAAAATATAAAGATCAATTTAT
<i>Bs4C-R</i>	121	CCCATGTTTGATCAGAGGGATTAATCACTGGTGAGATAACTTATCCCCCTATTACACCAT
<i>Bs4C-S</i>	121	CCCATGTTTGATGAGAGGGATTAATCACTGGTGAGATAACTTATCCCCCTATTACACCAT
<i>Bs4C-R</i>	181	AGTGATGAATAACTTATCCCATATTCACAGTGAGATAAGTTATCCCAAAAAATCCCAAT
<i>Bs4C-S</i>	181	AGTGATGAATAACTTATCCCATATTCATAGTGAGATAAGTTATCCCAAAAAATCCCAAT
<i>Bs4C-R</i>	241	ATTAAATTATCCAAAAATAAATTTCCCAACCAACGAACCCCTTTAAGTTAAGAGACTA
<i>Bs4C-S</i>	241	ATTAAATTATCCAAAAATAAATTTCCCAACCAACGAACCCCTTTAAGTTAAGATTTTC-
<i>Bs4C-R</i>	301	TAGGAAACGCAAGATTGACTTATGGAGAGTTCTAAGTCCGAATTGTTTTGTCAAATGG
<i>Bs4C-S</i>	296	-----CACAGATTGACTTATGGAGAGTTCTAAGTCCGACATTGTTTCGTCAAATGG
<i>Bs4C-R</i>	361	GGCTACGAAAAACATTATTTCTTCGCGAGGAACACAAGTAGCTTGACTACACTAACGTCC
<i>Bs4C-S</i>	347	GGCTACGAAAAACATTATTTCTTCGCGAGGA-CACAAGTAGTTTGACTACACTAACTTCC
<i>Bs4C-R</i>	421	CAGTGGAAAAGTACCTTCATGCAGCAGTATGGTACAATATTATTAAGCGCTGGAACATT
<i>Bs4C-S</i>	406	CAGTGGAAAAGTACCTTCATATAGCATATGGTACAATATTATTAAGCGCAGGAACATT
<i>Bs4C-R</i>	481	TTAAATaAAAAATAGTCCTCTCAGCCCTTAGACAAGCAGGAATAAAA-TTATTCTAAAT
<i>Bs4C-S</i>	466	TTAAATAcAAAAATAGTCCTCTCAGCCCTTAGACAAGTAGGATTAATAATTATTCTAAAT
<i>Bs4C-R</i>	540	CAACTTTCCGGATTGGGATCTTCGAGGAAAGAGCACCATCTTAAATGACAGTGTCTATA
<i>Bs4C-S</i>	526	CAACTTTCCGGATTGGGATCTTCGAGGAAAGAGCACCATCTTAAATG-----TA
<i>Bs4C-R</i>	600	AATAAATACACAACCAGCATACTACACCCAACCATAGAATTTTAGGAAAAGATTCAAAG
<i>Bs4C-S</i>	576	AATAAATACACAACCAGCATACTACACCCAACCATAGAATTTTAGGAAAAGATTCAAAG
<i>Bs4C-R</i>	660	AATAGA--ATGGAGTTTGATCTCAGATACTTGATCTTGATTTGGCTAACATGCTCAAAT
<i>Bs4C-S</i>	636	AATATAATATGGAGTTTGATCTCAGATACTTGATCTTGATTTGGCTAACATGCTCAAAT
<i>Bs4C-R</i>	718	CAATATTATC---CATTTCTGATAACTGGGATCCTTTCCATATATTTTCATGACCATCCCA
<i>Bs4C-S</i>	696	CAATATTATCATCCATTCTGATAACTGGGATCCTTTCCATATATTTTCATGACCATCCCA
<i>Bs4C-R</i>	775	GTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA
<i>Bs4C-S</i>	756	GTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA
<i>Bs4C-R</i>	835	TCACTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTACTTCAG
<i>Bs4C-S</i>	816	TCACTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTACTTCAG
<i>Bs4C-R</i>	895	CAGATCTCTCCAAGTCCTTTAATATCTTGTGCTAGCTTCTCTTCTACTCCACAAATGT
<i>Bs4C-S</i>	876	CAGATCTCTCCAAGTCCTTTAATATCTTGTGCTAGCTTCTCTTCTACTCCACAAATGT
<i>Bs4C-R</i>	955	TGTTCTGGTACTTTTCTCATTTATCATTGCACCTTTCCTCATGTTCTTCTTGGATTGGCA
<i>Bs4C-S</i>	936	TGTTCTGGTACTTTTCTCATCACCATTAACCTTTCCTCATGTTCTTCTTGGATTGGCA
<i>Bs4C-R</i>	1015	ATACATGGGCTAGTTTTCGACAAACGGATCTCGCATATTTTCTCAACAATAATATTTCCAG
<i>Bs4C-S</i>	996	ATCAGATGGGCTAGTTTTCGAGGAAGGACTCTGCATATTTTCTCAACAATA--TTTCCAG
<i>Bs4C-R</i>	1075	CAGTCAGCATTTTCATCAATGTGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAA
<i>Bs4C-S</i>	1053	CAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAA
<i>Bs4C-R</i>	1135	CTCAAACAAATCGCCTGCTCGTCAGTTGAA-----TTCATCGATCAAGGATATATTTCA
<i>Bs4C-S</i>	1113	CTCAAACAAATCGCCTGCTCGTCAGTTGAAATGAAATTCATCGATCAAGGATATATTTCA
<i>Bs4C-R</i>	1190	TGTTTCATCCAACGTCTTGTATTTTCATATATGTGTGTTTTTTTCTGTATTGAATTATGAAA
<i>Bs4C-S</i>	1173	TGTTTCATCCAACGTCTTGTATTTTCATATATGTGTGTTTTTTTCTGTATTGAATTATGAAA
<i>Bs4C-R</i>	1250	ATAAGGTACGTTTCCGTACTATAACAGGTAAAAGATGTCATGACAGTGTGAAATTTATGG
<i>Bs4C-S</i>	1233	ATAAGGTACGTTTCCGTACTATAACAGGTAAAAGATGTCATGACAGTGTGAAATTTATGG
<i>Bs4C-R</i>	1310	GTGCTATTACATGGACTACTCTTTTATACCAATATAGATaTAGTTGTTTAGCGATGTGT
<i>Bs4C-S</i>	1293	GTGCTATTACATGGACTACTCTTTTATACCAATATAGATCTATGTGTTTAGCGATGTGT
<i>Bs4C-R</i>	1370	ATACATAAATGGATACATGAGCTTTGTACCTTGATGGATATTCATCATCTGCCGAAGCA
<i>Bs4C-S</i>	1353	ATACATAAATGGATACATGAGCTTTGTACCTTGATGGATATTCATCATCTGCCGAAGCA
<i>Bs4C-R</i>	1430	TTTTCCATGTTAACTATTGTGTTGAATGCTGCTGATACTGTTTAATTAGTCACTTGTATG

Bs4C-S 1413 **TTTTCCATGTTAACTATTGTGTTGAATGCTGCTGATACTGTTTAATTAGTCACCTGTATC**

Bs4C-R 1490 **TTTGAGTAAAGCTTTTCTATAATATTCTTGTGTATTAC** **AAGTATTCTAAGTTGTACT**

Bs4C-S 1473 **TTTGAGTAAAGCTTTTCTATAATATTCTTGTGTATTAC** **AAGTATTCTAAGTTGTACT**

Bs4C-R 1550 **AATTTCAATTTCTCTTCTCTTCCTCCTTG**

Bs4C-S 1533 **AATTTCAATTTCTCTTCTCTTCCTCCTTG**

### Proteinsequenzen

>Bs4C-R

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSIFISITRITLHHPNIRIRVTRTTTSADLSKSFNIL  
CLASLLLPQMLFWYFFVIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>Bs4C-S

MEFDLRYFILILANMIKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSIFISITRITLHHPNIRIRIRTTTSADLSKSFNI  
LCLASLLLPQMLFWYFFFTITLSSCSSWIGNAWASFRGRTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

### Sequenzvergleich von Bs4C-R und Bs4C-S (93,9 % Identität)

Bs4C-R 1 **MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSIFISITR**

Bs4C-S 1 **MEFDLRYFILILANMIKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSIFISITR**

Bs4C-R 60 **ITLHHPNIRIRVTRTTTSADLSKSFNILCLASLLLPQMLFWYFFVIIALSSCSSWIGNTW**

Bs4C-S 61 **ITLHHPNIRIRIRTTTSADLSKSFNILCLASLLLPQMLFWYFFFTITLSSCSSWIGNAW**

Bs4C-R 120 **ASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV**

Bs4C-S 121 **ASFRGRTLHIFS-TIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV**

### A 7

Nukleotidsequenz des Kandidatengens 630 aus der *C. pubescens* Akzession PI 235047 und PI 585270

Vorhergesagte Start- und Stoppcodons sind dickgedruckt grün bzw. rot markiert. Der graue Hintergrund zeigt die kodierende Sequenz, wobei kleine Buchstaben Introns darstellen, die durch den Vergleich mit der cDNA des Kandidat 630 identifiziert wurden. Das *EBE<sub>AvrBs4</sub> Cand630* ist gelb hinterlegt.

>Kandidat 630 (PI 585270)

ACAGCTATATTTAAGTGCAATAATAATAGTACTGATTGGTAACCTATATTTAAGCATTGAAGTACTAGGAACTATGGGATT  
AAATTTGATAATCCTAAATTATTAAGGACTCTACTCTGTAATTTCCCTACACGTCGGCCTGTCCCTTTAGAGCGACACACCC**TATA**  
**AATAATAATCCCAAT**AAAAAGCAGCTCTCCATTTGGGAATTAACACAGTTGCTTGTCGATTGTTTGCACAAAA**ATGGCG**AAA  
GATTCTCCAAACCTCCCAAGGtaagtttgaataaaattaacatttggaaaccgagatctgggaacttcatgtagtctttatgaatca  
atttttggttattttttacctttacaatggctattttgtgttagcttttgattagagtatACGACACCAAGGGCGCCAACGAATCC  
AACAGTGAGCATTTGGGCAGCACAAATGTGGGAAATGCTTCAAATGGAGGACACTGTCAACACAGGAAGAGTTTGAGGAAATCAGGA  
GTAGGTTTGTCTGAACAACCTTTAACTGTGACAAACCAATATGGTTCTGTGTGATGACCCACCTGATATCGAGTACGATTCTTCT  
CGAACTTGGGCAATTGACAAACCTAACTCTCCCTAAAACCCATCTGGGTTTAAAGAGGGAGTTGTATCTCAGGAGAGATTACTCTAA  
AATGGATGCTTATTATGTCACTCCTACTGGCAAGAGACTAAGATCCCTTATTGAAGTGGGTAATTTTCTTCAAAATAATCCTGAGT  
TCAGGGACTTATCAGTCTCTGATTTCAGTTTGTGAGCCCCAAAATATGGATGATACTATACCCGCTACTGCTGTTGTTGCCAAC  
TCTAACAAGAAGGGTGCACCTAGTACTGCTAAA**TGA**GTGCGTGGATGGAGCTCGAATTGTTTGTGTTTGTATCTGTAGATGCAGTATT  
TAGAGTCCCTGGTATGCTTGAAGGTTAAACTTATTTCCCGATGTATGTTATTTTCAATGTCTTTGGTATGCAAATCTCTTATGTATGT  
GGAAACATCTCAACTGCTTGGTTACCTTTTGTGCGGAAGGATTGATATCAAAGATTTCGAACCTATTTTCTAGGGGTGCGTTAC  
GACTTTAGTGTTCAGATTCTGTACATAAAGTTGTGGTTGTGGTGTGTAATTTCCCTAGTTTGTGCGGTGTTTCTATGAATGAT  
CCTTTCCTGGTTCA

>Kandidat 630 (PI 235047)

ACAGCTATATTTAAGTGCAATAATAATAGTACTGATTGGTAACCTATATTTAAGCATTGAAGTACTAGGAACTATGGGATT  
AAATTTGATAATCCTAAATTATTAAGGACTCTACTCTGTAATTTCCCTACACGTCGGCCTGTCCCTTTAGAGCGACACACCC**TATA**  
**AATAATAATCCCAAT**AAAAAGCAGCTCTCCATTTGGGAATTAACACAGTTGCTTGTCGATTGTTTGCACAAAA**ATGGCG**AAA  
GATTCTCCAAACCTCCCAAGGtaagtttgaataaaattaacatttggaaaccgagatctgggaacttcatgtagtctttatgaatca  
atttttggttattttttacctttacaatggctattttgtgttagcttttgattagagtatACGACACCAAGGGCGCCAACGAATCC  
AACAGTGAGCATTTGGGCAGCACAAATGTGGGAAATGCTTCAAATGGAGGACACTGTCAACACAGGAAGAGTTTGAGGAAATCAGGA  
GTAGGTTTGTCTGAACAACCTTTAACTGTGACAAACCAATATGGTTCTGTGTGATGACCCACCTGATATCGAGTACGATTCTTCT  
CGAACTTGGGCAATTGACAAACCTAACTCTCCCTAAAACCCATCTGGGTTTAAAGAGGGAGTTGTATCTCAGGAGAGATTACTCTAA  
AATGGATGCTTATTATGTCACTCCTACTGGCAAGAGACTAAGATCCCTTATTGAAGTGGGTAATTTTCTTCAAAATAATCCTGAGT  
TCAGGGACTTATCAGTCTCTGATTTCAGTTTGTGAGCCCCAAAATATGGATGATACTATACCCGCTACTGCTGTTGTTGCCAAC  
TCTAACAAGAAGGGTGCACCTAGTACTGCTAAA**TGA**GTGCGTGGATGGAGCTCGAATTGTTTGTGTTTGTATCTGTAGATGCAGTATT  
TAGAGTCCCTGGTATGCTTGAAGGTTAAACTTATTTCCCGATGTATGTTATTTTCAATGTCTTTGGTATGCAAATCTCTTATGTATGT  
GGAAACATCTCAACTGCTTGGTTACCTTTTGTGCGGAAGGATTGATATCAAAGATTTCGAACCTATTTTCTAGGGGTGCGTTAC  
GACTTTAGTGTTCAGATTCTGTACATAAAGTTGTGGTTGTGGTGTGTAATTTCCCTAGTTTGTGCGGTGTTTCTATGAATGAT  
CCTTTCCTGGTTCA

## A 8

Sequenzalignment der Nukleotidsequenzen des *Kandidat 630* aus den *C. pubescens*-Akzessionen PI 235047 und PI 585270.

Identische Nukleotide zwischen beiden Akzessionen sind als weiße Buchstaben auf schwarzem Hintergrund dargestellt. Vorhergesagte Start- und Stoppkodons sind fett grün bzw. rot markiert. Kleine Buchstaben geben Introns an, die durch den Vergleich mit der cDNA des *Kandidat 630* identifiziert wurden. Das *EBE<sub>AvrBs4</sub>Cand630* ist gelb hinterlegt. Das Alignment wurde mit ClustalW und Boxshade erstellt.

PI585270	1	ACAGCTATATTTAAGTGCAATAATAATAGTACTGATTGGTAACCTATATTTTAAGCATT
PI235047	1	ACAGCTATATTTAAGTGCAATAATAATAGTACTGATTGGTAACCTATATTTTAAGCATT
PI585270	61	GAAGTAGGCTAGGAACTATGGGATTAAATTTGATAATCCTAAATTATTAAGGACTCTAC
PI235047	61	GAAGTAGGCTAGGAACTATGGGATTAAATTTGATAATCCTAAATTATTAAGGACTCTAC
PI585270	121	TCTGTAATTTCCCTACACGTCGGCCTGTCCCTTTAGAGCGACACACCC
PI235047	121	TCTGTAATTTCCCTACACGTCGGCCTGTCCCTTTAGAGCGACACACCC
PI585270	181	TCCCAATAAAAAGCAGCTCTCCATTTTGGGAATTAACACAGTTGCTTGTCGATTGTTT
PI235047	181	TCCCAATAAAAAGCAGCTCTCCATTTTGGGAATTAACACAGTTGCTTGTCGATTGTTT
PI585270	241	GCACAAAAAATGGCGGAAAGATTCTCCAAAACCTCCCAAGGtaagtttgaataaaattaac
PI235047	241	GCACAAAAAATGGCGGAAAGATTCTCCAAAACCTCCCAAGGtaagtttgaataaaattaac
PI585270	301	atttgaaccgagatctgggacttcatgtagtctttatgaatcaatttttggttattttt
PI235047	301	atttgaaccgagatctgggacttcatgtagtctttatgaatcaatttttggttattttt
PI585270	361	tacctttacaatggctattttgtgttagcttttgattagagtatACGACACCAAGGGCGC
PI235047	361	tacctttacaatggctattttgtgttagcttttgattagagtatACGACACCAAGGGCGC
PI585270	421	CAACGAATCCAACAGTGAGCATTGGGCAGCACAATGTGGGAAATGCTTCAAAATGGAGGA
PI235047	421	CAACGAATCCAACAGTGAGCATTGGGCAGCACAATGTGGGAAATGCTTCAAAATGGAGGA
PI585270	481	CAGTGTCAACACAGGAAGAGTTTGAGGAAATCAGGAGTAGGTTTGCTGAACAACCCCTTA
PI235047	481	CAGTGTCAACACAGGAAGAGTTTGAGGAAATCAGGAGTAGGTTTGCTGAACAACCCCTTA
PI585270	541	ACTGTGACAACAAACCTAATGGTTCTTGTGATGACCCACCTGATATCGAGTACGATTCTT
PI235047	541	ACTGTGACAACAAACCTAATGGTTCTTGTGATGACCCACCTGATATCGAGTACGATTCTT
PI585270	601	CTCGAACTTGGGCAATTGACAAACCTAATCTCCCTAAAACCCCATCTGGGTTTAAGAGGG
PI235047	601	CTCGAACTTGGGCAATTGACAAACCTAATCTCCCTAAAACCCCATCTGGGTTTAAGAGGG
PI585270	661	AGTTGTATCTCAGGAGAGATTACTCTAAAATGGATGCTTATTATGTCACTCCTACTGGCA
PI235047	661	AGTTGTATCTCAGGAGAGATTACTCTAAAATGGATGCTTATTATGTCACTCCTACTGGCA
PI585270	721	AGAGACTAAGATCCCTTATTGAAGTGGGTAATTTTCTTCAAAATAATCCTGAGTTCAGGG
PI235047	721	AGAGACTAAGATCCCTTATTGAAGTGGGTAATTTTCTTCAAAATAATCCTGAGTTCAGGG
PI585270	781	ACTTATCAGTCTCTGATTTTCAAGTTTGTGAGCCCCAAAATTATGGATGATACTATACCCG
PI235047	781	ACTTATCAGTCTCTGATTTTCAAGTTTGTGAGCCCCAAAATTATGGATGATACTATACCCG
PI585270	841	CTACTGCTGTTGTTGCCAACTCTAACAAGAAGGGTGCACCTAGTACTGCTAAA
PI235047	841	CTACTGCTGTTGTTGCCAACTCTAACAAGAAGGGTGCACCTAGTACTGCTAAA
PI585270	901	GTGGATGGAGCTCGAATTGTTTGTGTTTGTATCTGTAGATGCAGTATTTAGAGTCCTGGTAT
PI235047	901	GTGGATGGAGCTCGAATTGTTTGTGTTTGTATCTGTAGATGCAGTATTTAGAGTCCTGGTAT
PI585270	961	GCTTGAAGGTTAACTTATTCCCGATGTATGTTATTTTCAATGTCTTTGGTATGCAAATC
PI235047	961	GCTTGAAGGTTAACTTATTCCCGATGTATGTTATTTTCAATGTCTTTGGTATGCAAATC
PI585270	1021	TCTTATGTATGTGGAAAACATCTCAACTGCTTGGTTACCTTTTGTGCGGAAGGATTTGA
PI235047	1021	TCTTATGTATGTGGAAAACATCTCAACTGCTTGGTTACCTTTTGTGCGGAAGGATTTGA
PI585270	1081	TATCAAAGATTGAACTCTATTTCTAGGGGTCGCTTACGACTTTAGTGTTCAGATTCTG
PI235047	1081	TATCAAAGATTGAACTCTATTTCTAGGGGTCGCTTACGACTTTAGTGTTCAGATTCTG

```

PI585270 1141 TACATAAAGTTGTGGTTGTGGTGCTGAATTTCCCCTAGTTTTTGTTCGGTGTCTCTATGAA
PI235047 1141 TACATAAAGTTGTGGTTGTGGTGCTGAATTTCCCCTAGTTTTTGTTCGGTGTCTCTATGAA

PI585270 1201 TGATCCTTTCCTGGTTCA
PI235047 1201 TGATCCTTTCCTGGTTCA

```

Proteinsequenz des Kandidat 630 aus den *C. pubescens*-Akzessionen PI 235047 und PI 585270.

> Kandidat 630 (PI 235047 und PI 585270)

```

MAKDSFKPPKTTTPRAPTNPTVSIWAAQCGKCFKWRTLSTQEEFEEIRSFAEQPFNCNDKPNQSCDDPPDIEYDSSRTWAIDKPNL
PKTPSGFKRELYLRDYSKMDAYYVPTPTGKRLRLSLIEVGNFLQNNPEFRDLSVSDFSFVSPKIMDDTIPATAVVANSNKKGAPSTA
K-

```

## A 9

Protein-Sequenzvergleich des *C. pubescens* Kandidat 630 mit einem Methyl-CpG-bindenden Protein aus Soja (*Glycine max*)

Proteinsequenz des vorhergesagten methyl-CpG-bindende Domäne-beinhaltenes Protein 4-like Isoform X1 [*Glycine max*]

NCBI Reference Sequence: XP\_003536387.1

>G.max

```

MVKYGREKREVTYVNNMKGQEI SKTPSSSKRTLSQGSVDIYAAQCKNCLKWREIDTQEEFEEIRSKVAEEPFLLCSRKANSSCDEP
GDIKYDSSRTWVIDKPNLPKTPQGFKRSLVLRKDYSKLDAYYITPAGKKLRTRNEIAAFLKDNPEFKGVASDFDFSSPKIMQDTI
PEVVEQKDSANKKVIAKDEV

```

Vergleich der Aminosäuresequenzen des Kandidat 630 (*C. pubescens*) mit einem Methyl-CpG-bindenden Protein aus Soja (*Glycine max*, G.max). Identische Nukleotide sind als weiße Buchstaben auf schwarzem Hintergrund dargestellt. Das Alignment wurde mit ClustalW und Boxshade erstellt. Die Identität beider Proteine beträgt 56%.

```

G.max.      1  MVKYGREKREVTYVNNMKGQEI SKTPSSSKRTLSQGSVDIYAAQCKNCLKWREIDTQEE
C.pub.      1  M-----AKDSPKPPKTTTPRAPTNPTVSIWAAQCGKCFKWRTLSTQEE

G.max.     61  FEEIRSKVAEEPFLLCSRKANSSCDEPGDIKYDSSRTWVIDKPNLPKTPQGFKRSLVLRKD
C.pub.     43  FEEIRSRFAEQPFNCNDKPNQSCDDPPDIEYDSSRTWAIDKPNLPKTPSGFKRELYLRD

G.max.    121  YSKLDAYYITPAGKKLRTRNEIAAFLKDNPEFKGVASDFDFSSPKIMQDTIPEVVEQKD
C.pub.    103  YSKMDAYYVPTPTGKRLRLSLIEVGNFLQNNPEFRDLSVSDFSFVSPKIMDDTIPATAVVAN

G.max.    181  SANKKVIAKDEV
C.pub.    163  SNKKGAPSTAK--

```

Vergleich der Sequenzen des Kandidat 630 (*C. pubescens*) mit einem Methyl-CpG-bindenden Protein aus *Arabidopsis thaliana* (AtMBD4, AT3G63030). Identische Nukleotide sind als weiße Buchstaben auf schwarzem Hintergrund dargestellt. Das Alignment wurde mit ClustalW und Boxshade erstellt. Die Identität beider Proteine beträgt 50%.

```

630      1  MAKDS--PKPPKTTTPRAPTN--TVSIWAAQCGKCFKWRTLSTQEEFEEIRSFAEQPFN
AtMBD4   1  MKEEEIIGKPAKPKAKKDVA PGRLIDTYAAQCDNCHKWRVIDSQEEYEDIRSKMLEDPFN

630     57  CDNKPNQSCDDPPDIEYDSSRTWAIDKPNLPKTPSGFKRELYLRDYSKMDAYYVPTPTGK
AtMBD4   61  CQKKQGMSCEEPADIDYDSSRTWVIDKPGLPKTPKGFKRSLVLRKDYSKMDTYFYPTPTGK

630    117  RLRLSLIEVGNFLQNNPEFRDLSVSDFSFVSPKIMDDTIP-----ATAVVANSNKKGAPST
AtMBD4   121  KLRSLRNEIAAEVVEANPEFRNAPLGDFNFTVPKVMEDTVPPDPKLGSPFPSTTTTSEKSS

630    172  AK----
AtMBD4   181  VKQSHN

```

## A 10

Nukleotid-Sequenzalignment und Vergleich des *C. pubescens* Kandidat 630 mit einem Methyl-CpG-bindenden Protein aus Kartoffel (*Solanum tuberosum*, SGN-U284104). Identische Nukleotide zwischen beiden Akzessionen sind als weiße Buchstaben auf

schwarzem Hintergrund dargestellt. Das Alignment wurde mit ClustalW und Boxshade erstellt. [http://solgenomics.net/search/unigene.pl?unigene\\_id=284104](http://solgenomics.net/search/unigene.pl?unigene_id=284104)

```
>S. tub. (Solanum tuberosum, SGN-U284104)
ATGGCGAAAGATTCTCCAAAACCTCCCAAGACGTCGTCGAGGTCGCCAACGAATCCATCAGTGAGCGTGTGGGCAGCACAAATG
TGAAAAATGCTTGAATGGAGGAGGATTGCGACACAGGAAGAATTTGAGGAAATCAGAAGTAGGTTTACTGAAGAACCCTTTAACT
GCCACAACCAACCAATAGTTCTTGTGATGACCCACAGATATTGAGTACGATTCTTCTCGAACTTGGGCTATTGACAAGCCCAAT
CTGCCTAAAACCCCATCTGGGTTTAAGAGAGAATTGTACCTGAGGAGAGATTACTCTAAGATGGATGCTTATTATTCACTCCTTC
AGGGAAGAACTCAGATCCTTTACTGAAGTAACCTACTTTCTTCAACAAAATCCCCAGTTCAGTGACGTTAAACCCCTCAGATTTC
GCTTTACCAGCCCCAAAGTTATGATTGATACCATACCCTCTACCCTCTTCTTGCCAACTCTCACAGAAGGGTGCAGCTAATAGC
ACTAAATGA
```

Sequences (1:2) Aligned. Score: 84.10

Guide tree file created: [clustalw2-I20140818-123138-0250-73397727-pg.dnd]

```
S. tub.      1  ATGGCGAAAGATTCTCCAAAACCTCCCAAGACGTCGTCGAGGTCGCCAACGAATCCA
C. pub.     1  ATGGCGAAAGATTCTCCAAAACCTCCCAAGACGACAC---CAAGGGCGCAACGAATCCA

S. tub.     61  TCAGTGAGCGTGTGGGCAGCACAAATGTGAAAAATGCTTGAATGGAGGAGGATTGCGACA
C. pub.     58  ACAGTGAGCATTTGGGCAGCACAAATGTGGGAATGCTTCAATGGAGGACACTGTCTACACA

S. tub.     121  CAGGAAGAATTTGAGGAAATCAGAAGTAGGTTTACTGAAGAACCCCTTTAACTGTCACAAAC
C. pub.     118  CAGGAAGAATTTGAGGAAATCAGGAGTAGGTTTGTGAAACAAACCCCTTTAACTGTGACAAAC

S. tub.     181  CAACCCCAATAGTTCTTGTGATGACCCACACAGATATTGAGTACGATTCTTCTCGAACTTGG
C. pub.     178  AAACCTAATAGTTCTTGTGATGACCCACCTGATATCGAGTACGATTCTTCTCGAACTTGG

S. tub.     241  GCTATTGACAAGCCCAATCTGCCTAAAACCCCATCTGGGTTTAAGAGAGAATTGTACCTG
C. pub.     238  GCAATTGACAACCTAATCTCCCTAAAACCCCATCTGGGTTTAAGAGGGAGTTGTATCTC

S. tub.     301  AGGAGAGATTACTCTAAGATGGATGCTTATTATTCACTCCTTCAGGCAAGAAACTCAGA
C. pub.     298  AGGAGAGATTACTCTAATATGGATGCTTATTATTCTACTCCTACTGGCAAGAGACTAAGA

S. tub.     361  TCCTTTACTGAAGTAACCTACTTTTCTTCAACAAATCCCAAGTTCACTGACGTTAAACCC
C. pub.     358  TCCCTTATTGAAGTGGGTAATTTTCTTCAAAATAATCCTGAGTTCAGGGACTTATCACTC

S. tub.     421  TCAGATTTCAGCTTTTACAGCCCCAAAGTTATGATTGATACCATACCCTCTACCGCTCTT
C. pub.     418  TCTGATTTCAGTTTCTCAGCCCCAAAGTTATGATGATGATACATATACCGCTACTGCTCTT

S. tub.     481  CTTGCCAACTCTCACAGAAGGGTGCAGCTAATAGCACTAAATGA
C. pub.     478  GTTGCCAACTCTAACAGAAGGGTGCACCTAGTACTCTAAATGA
```

## A 11

Nukleotidsequenz des Kandidatengens 15502 aus der *C. annuum* cv. CM334

Vorhergesagte Start- und Stoppcodons sind fett grün bzw. rot markiert. Der graue Hintergrund zeigt die kodierende Sequenz, wobei kleine Buchstaben Introns darstellen, die mit Hilfe des Vorhersageprogramms Softberry identifiziert wurden. Das *EBE<sub>AvrBs4</sub>* 15502 ist gelb hinterlegt.

```
>Kandidat 15502 (C. annuum CM334)
AGTTAAATAAAGTATTATTAAATATTTTTTTTTTTTTTTTATTCTTATCAAAGATTTTTTAAATAAAAAAATATTTAACAAT
CAAAAGCATTCATAATAAGCAAAATTCAAACACCCACGTCATTAATATATTTACACTTTGCAGTAAGAGACAAGATAATAGTCCAT
GATTTTATTATCTTTGTGAAGAGCACATGGACGGTCATTTCTCTGCCTCGGCCACAAAGCAATTGACGATCAGAATTTTTTCGTCC
GTCCGCACGCATCTCCCCAAAAAATACATCAAGTGGTATCGATTTCGATACATCCACCATCAGTAACCTTTTTTCTCCTCAGAA
ATCTCGGGATGAGTGAATGTGGTATGATCAATAATCCACATCTCCACATTCTCTCACCTTCATAAATCTCCAATTTTCACTCTC
ACACTCCACTCACATTTCAATTGTTGAAAACTAGATCTGTGCATGAGCATAATGGCGACATCGTCTACTGTGCCCGCTACAATGG
AGTAGAGTTATTCAACCATACGGGTCCAACCTCCAGATTGATCCGTTTCAGTTCTATAAACACAAGTTCCAATTTGTTCTTACTA
GAACCTCCAATTTTCGAGAGCCAAACGTTGTTTCTATGTGAAGAGTACCCTTAAGTGAGGCCAAGCAGAAAGTTTCAGGATCCTATT
ACTGACCTAGTgattttccttaatttaatttctgtcaatttatttggaatcaaaattgggtaacatttcgtgggtcaatcatttagga
aaggcataacacataaatgtgcccttcaacttggccttaactggctctcgaaactttggatttgcaacacacagcctcctatgtagc
aaactcgtgatttaacatgttaggaagcgtgcacctcataggattttccaaattaagtttcttgatagtttggacctgccaatctgg
ttattgatcaaatcttatgttggtaaggcattattgcagtatcctaactcgttcttattcattttgtatacaatagtgaaagcattaga
aacttggatccggaccataggaattacagaaacttttgatataatacataagccctttaaacttgggcatctatgccctccaattt
tggatgtgcacaagtagacaagttgaacatgtaaacacatatgttctacgtggaaaatctcatgattacatgtaggacgcatgca
tctaacacattttgccataggaatgtgtgtacttgttctacttgaatctatctcttgatgagataagctcttttactttatata
atgaacatcaggatttagctgtgttcttctaaatcaaaataagcagagtcagtggttagggttttggtaaatgaaaaacttatagaga
gatttacctgtttgcaccgaatttatcctcctgtgtgttgcacaaaattttcaaaatattttgccttttgaattacaacattaa
cataatatgactagaccatcattttacaataatatgcaagaattgatgggttatttcaagataaatgatggaacttcgaaggat
```

[illegible]

**GTTTAGCAGTGACAGAACAATCCACGAATATGCCAAAGACATTTGGAACGTTGAACCTGTGGAGTTACCATAAGAGAGGGAAGGGA**  
 ATAAAAACAACAAAGCCTCAGTGAGTAGCATTGACTAGTTCACCTTGAGTCGTTTCTTTTTTTAGTGTGAAGGTATATAATCTGTGA  
 AGAGGAATGATGTTTACCACATAATAAGAGATCAAAGAGCCTCATTTGCTGCTCAGTTGATCTCTTAATTTATGTTTGGTGCAC  
 ATACTCAAACCCGGCAGATCCTCTGGGTGGTGTATGTCATTAATGTTTATGCTGCCTAATCCATCAAAATAGATGATTTTAAATCA  
 AGTTCATGCTTTTATTCCTAAAATAAAATGGAGGTGAGAGATTAGGGACAAGAAAGTAGTGGAAGATTGTGCAACCGCCAAG  
 GGTGTCATGATGGATGTAGTGCTGACAGAATAGATAAAATTTGTGACAGATATTCGTCGAATAAAAAATGGATTGGATAATTAATTTAC  
 TCAAGTCTCAACAAATGTGGATCGCTCCATTCTATTTCGCTGTAAAAATGTATTACACTAGCTATCGGAATCGTACTTCCATTCTAT  
 CTTTCTTTTTTTAAAGATTGTTTTTGCCTTGAGCTGAGGGTTCATTGAGAACACCTCTTACCTCATTTTGGAGATAGAGGTAAAGT  
 TTCCTTCTACACTTCTAACTCCATTTTGTGGGACTATAATGTTATTTTCGGATACAGAATCTGTATATAACAATTTGAACCTCA  
 AACTAAGTTCAAGAACAACCTTTAAAACTACAATAAAGTTCAACAATTTCAAACATAAGTTCAAATGAGTGATCAAATGAAGTAC  
 TAAAATGAAGTTTCTCAAGTGAATGATCAAATGAAGTATCAAATAAAGTTTCTACTATTTTAGAAAAATAGTAAACCAACTTA  
 AGAAATGAAGATGAAACAGAAAAAAGGATAGAATTAAAGTCCATCGAATTCACGGCGTGTGCTTAAGGAATTTATTCTTCTCAAA

## A 12

Promoter- und cDNA-Sequenz des Kandidatengens 15502 aus der *C. pubescens* Akzession PI 235047 und PI 585270

Vorhergesagte Start- und Stoppcodons sind fett grün bzw. rot markiert. Das *EBE<sub>AvrBs4</sub>Cand15502* ist gelb hinterlegt.

>Kandidat 15502 (PI 235047 und PI 585270)

CACTTTGTAGTAAGAGACAAGATAATAGTCCATGATTTTATTATCTTTGTGAAGAGCACGTTGGACGGTCATTCTCTGCCTCGGCC  
 CACAAAGCAATTAACGATACCATTTTTTCGTCCGTCGCTTCTCCCCAAAAAATCACATCAAGTGGTATCGATTTCGATACATCC  
 ACCATCAGTAAGTGTTTTTTCTCAGAAATCTCGGGATGAATGAAATGTGATATGATCAATAATGTTTAAAACTAGATCTGCGCCTGAGCATAATGGCG  
 ACATCGCTCTACTGTGCCACTACAAATGGAGTAGAGTTATTAACCATACGCGGTCCAAGTCCAGATGATCCGTTGATTCATCTCTAT  
 AAACACAAGTTTCCAATCGTTTTCTACTAGAACCTCCAATTTTCGAGAGCCAAACGTTGTTTCTATGTGAGAGTACCTTAAGTG  
 AGCCCAAGCAGAAAGTTCAGGATCCTATTACTGACCTAGGCGCTGAGAGTGATCTGAGTTCTTTTGTCTCTGATGCTGCATCTATT  
 GCCTCAAGTATCAATACCATGCAGAATTACACCTTTATATTCTCTGAAAGGTTTGAGTCCCGAAGGCATTCTTTGCAACAGC  
 CCAAAGTGTCGGTGATTCCCTCCTTATTAATTGGAATGCTACGTATGATACTTACGAAAAGAAGACATGAAGCAAGCATACTATC  
 TCTCCATGGAATTTCTTCAGGGTAGAGCATTTGTTGAACGCAATTTGGTAATCTGGAGCTTACTGGTGCTTATGCGGAAGCTTTAAAA  
 AAGCTTGGCCACAATCTAGAAAATGTGGCTTCTCAAGAGCCAGATGCTGCTCTTGGAAATGGAGGTTTGGGAAGGCTTGCTTCCTG  
 TTTTCTGGACTCTTTGGCGACACTAACTACCCAGCATGGGGCTATGGACTTAGGTACAAGTATGGTTTGTAAAGCAACGGATTA  
 CAAAAGATGGTCAGGAAGAGATCGCTGAAGATTGGCTTGAATTTGGCAGTCCATGGGAAGTTGTGAGGAATGATGTTTCATATCTT  
 ATCAAAATTTCTATGGAAAGTCACTACAGGATCAGATGGAAAGAGATATTGGATTGGTGGAGAGGATATAAAGGCAGTTGCGTATGA  
 TGTTCCCATTTCCAGGGTATAAAACCAACACGACAATCAACCTTCGACTGTGGTCTACACAGGTTCCATCAGCGGACTTTTATTTAT  
 CTGCTTTCAATGCTGGAGAGCACACCAAGCATGTGAAGTCGAAGCAACGCCAAGAAGATATGTTACGTAATCTACCTCGGGGAT  
 GAATCAGATGATGGAAAAATCCTTAGGTTGAAGCAACAATATACCTTGCTGCTCGGCTTCTCTCAGGATATTATTGCACGATTGGA  
 GAGGATCATGTGATCGTATTAGTGGGAAGAGTTTCTCGGAAAAAGTTGCTGTGCAGATGAATGACACTACCCCTACCTTTTGTA  
 TCCCTGAGCTGATGAGAATATTATAGATCTGAAGGCTTGAATTGGAATGAAGCTTGGAAATATACTAAAAAGAACTGTGGCATA  
 ACAAAACCATACTGTTTTGCCTGAGGCATTGGAGAAATGGAGTTATGAATTGATGCAGAAGCTGCTTCCAGACATGTTGAAATCAT  
 TGAGGCGATTGACGAGGAGCTGGTACAAGAAATGTATCGAAATTTGGCTCACTGGATCTGGAAAAATTGGAGGAAAAAGTTGACTA  
 CGATTGAAATTTTAGAAAAATTTTGATCTTCCAGTTCTGTGCTGATTTTACTAAGCCTGAAATCCAGTTGATAATAATACT  
 GAAACAGTAGAAGTCCATGACAAAGTTAAAGCTTCTGGTGAAGTTGTAACATATGATGAAGATGACACTGGTAAGAAAACTAGTGC  
 AAAGACAGAAAGCAGCTGCAGAAAAAGACATTGAAACGAAAACTATTGTGAGATTGGAACAGCTGTCATACCTTCTAAGAAGGTTTC  
 GCATGGCCAACTTGTGTGTTGTTGGTGGTCAATGCTGTTAATGGCGTCGCTGAGATCCATAGTGAATTTGTGAAGGAGGAGGTTTTC  
 AGAGACTTCTATGAGCTCTGGCCTGAAAGTTTCCAAAGCAACAAACAAATGGAGTTACACCAAGAAGATGGATTCAATTTCTGCAATCC  
 TGCTCTTAGTAGCATCATACTAAATGGACCGGTACAGAGGACTGGGTCTGAAACTGAAAGTTAGCAGAATTACAGAAGTTTG  
 CTGATAATGAAGTTCTTCAAACCTGAGTGGAGGGAAGCAAAAGGAGCAACAGATTAAAGTTGCCTCCTTTCTCAAGAAAAAGACA  
 GGGTATTCTGTTGTCCAGATGCAATGTTTGATATTGAGTAAACGCATTACAGAGTACAAGCGACAATCTAAATATCTTTCGG  
 GATTGTTTATCGGTATAAGAAAGATGAAAGAAATGACAGCTGCAGAAAGAAAGTCTAATTTGCTTCTCGGTGCTCAAGCTCATGAAATTCAGGGCT  
 GAAAGCTTTTGCCACATATGTGCAAGCCAAAGAGGATTGTAATAATTTATCAGATGTTGGGGCTACTATAAATCATGATCCAGAA  
 ATAGGTGATCTGTTGAAGGTAGTATTTGTGCCAGATTATAATGTTAGTGTGCTGAATTGCTAATTCCTGCTAGTGATCTATCACA  
 ACATATCAGTACGGCTGGAATGGAGGCCAGTGAACCAAGTATGAAGTTTGAATGAATGGTTGTATCCAAATTTGGTACATTTGG  
 ATGGTGCTAACGTGGAATAAGACAAAGAGTTGGAGAAGAGAACTCTTTCTCTCGGTGCTCAAGCTCATGAAATTCAGGGCTT  
 AGAAAGAAAGAGCTGACGGAAAGTTTGTGCTGATGAACGTTTGAAGAGGTCAAGGAATTTGTTAGAAGCAGTGTTTTGGCTC  
 TTATAACTATGATGAGCTAATTGGATCCTTGGAAAGGAAATGAAGGTTTTTGGTCGTGCTGACTATTTCTTGGGGCTGAGACTTCC  
 CCAGTTACATAGAATGCCAAGAGAAAGTTGATGAGGCATATCGCGACCAAAAAAGGTGGACAAAGATGTCAATCTTGAATACAGCA  
 GGATCTACAAATTTAGCATGTGACAGAACATCCACGAATTTGCCAAAGACATTTGGAACATTGAACCTGTGAAGTTACCA**TAA**

## A 13

Proteinsequenz des Kandidatengens 15502 aus der *C. pubescens* Akzession PI 235047 und PI 585270

>Kandidat 15502 (PI 235047 und PI 585270)

MATSSYCATTNGVELFNHYGSNSRLIRFSSINTSSKSFLTRTSNFRRAKRCFYVKSTLSEPKQKVQDPITDLGAESDLSSFAPDAA  
 SIASSIKYHAEFTPLYSRPERFELPKAFFATAQSVRDSLLNWNATYDYTYEKKNMKQAYYLSMEFLQGRALLNAIGNLELTGAYAEA  
 LKKLGHNLNVASQEPDAALNGGLGRASCFLDSLATLNPWAGYGLRYKYLFLKQIRITKDGQEEIAEDWLEIGSPWEVVRNDVS  
 YPIKFYGVKVTGSDGKRYIWGGEDIKAVAYDVPIPGYKNTNTINLRLLSTQVPSADFYLSAFNAGEHTKACEVEANAKKICYVLYP  
 GDESDDGKILRLKQYTLCSASLQDIARFERRSCDRIKWEFPEKQVAVQMNDRHTPLCIPELMRIFIDLKGLNWNENWITKRTV  
 AYTNTHTVLPEALEKWSYELMQKLLPRHVEIEAIDEELVQEIIVSKFGLDLEKLEKLTMRILENFDLPSSVADLFTKPEIPVDN  
 NTETVEVHDKVKASGEVVTNDEDDTGKTSKTEAAAEKDIETKTIVRLEPAVILPKKVRMANLVCVGGHAVNGVAEIHSEIVKEE

VFRDFYELWPEKFQNKTNQVTPRRWIHFCNPALSSIIITKWTGTEDWVLKTEKLAELQKFADNEVLQTEWREAKRSNKIKVASFLKE  
KTGYSVVPDAMFDIQVKRIHEYKRQLLNIFGIVYRYKKMKEMTAERKSNFVPRVCIFGGKAFATYVQAKRIVKFITDVGATINH  
PEIGDLLKVVFPDYNVSVAEELLIPASDLSQHISTAGMEASGTSNMKFAMNGCIQIGTLDGANVEIRQEVGEENFFLFGAQAHEIA  
GLRKERADGKFVPDERFEEVKEFVRSSVFGSYNYDELIGSLEGNEGFGGRADYFLVGMDFPSYIECQEKVDEAYRDQKRWTMSILN  
TAGSYKFSSDRTIHEYAKDIWNIEPVKLP

#### A 14

Proteinsequenz der alpha-glucan Phosphorylase aus *Solanum tuberosum*  
alpha-1,4 glucan phosphorylase L-1 isozyme, chloroplastic/amyloplastic-like [*Solanum tuberosum*] (Sequence ID: [ref|NP\\_001275215.1](#))

#### > alpha-glucan Phosphorylase (*S. tuberosum*)

MATANGAHLFNHYSSNSRFIHFTSRNTSSKLFLLTKTSHFRPKRCFHVNNLTSEKIHHPITEQGGESDLSSFPADDAASITSSIKYH  
AEFTPVFSPERFELPKAFFATAQSVRDSLLINWNATYDIYEKLNMQAYYLSMEFLQGRALLNAIGNLELTGAFAEALKNLGHNLE  
NVAEQEPDAALGNGGLGRLASCFLDSLATLNYPAGYGLRYKYGLFKQRITKDGQEEVAEDWLEIGSPWEVVRNDVSYPIKFYGV  
STGSDGKRYWIGGEDIKAVAYDVPIPGYKTRTTISLRLWSTQVPSADFDLSAFNAGEHTKACEAQANAEEKICYILYPGESEEGKI  
LRLKQQYTLCASLQDIISRFEERRSGDRIKWEFPEKVAVQMNDHTPTLCIPELMRILIDLKGLNWEAWNITQRTVAYTNHTVLP  
EAEKWSYELMQKLLPRHVEIEAIDEELVHEIVLYKYGMDLNKLEELTMRILENFDLPSSVAELFIKPEISVDDDTETVEVHD  
KVEASDKVVTNDEDDTGKKTSVKIEAAAEKDIKKTTPVSEPAVILPKKVRMANLCVVGHHAVNGVAEIHSEIVKEEVNDFYELW  
PEKFQNKTNQVTPRRWIRFCNPPLSAIIITKWTGTEDWVLKTEKLAELQKFADNEVLQNEWREAKRSNKIKVVSFLKEKTGYSVVPD  
AMFDIQVKRIHEYKRQLLNIFGIVYRYKKMKEMTAERKTNFVPRVCIFGGKAFATYVQAKRIVKFITDVGATINHDP  
EIGDLLKVVFPDYNVSVAEELLIPASDLSQHISTAGMEASGTSNMKFAMNGCIQIGTLDGANVEIRQEVGEENFFLFGAQAHEIA  
GLRKERADGKFVPDERFEEVKEFVRSGAFSGSYNYDDIGSLEGNEGFGGRADYFLVGKDFPSYIECQEKVDEAYRDQKRWTMSILN  
TAGSYKFSSDRTIHEYAKDIWNIEAVEIA

Protein-Sequenzalignment des *C. pubescens* Kandidat 15502 mit einer alpha-Glucan Phosphorylase aus Kartoffel (*Solanum tuberosum*)

Weißer Buchstaben auf schwarzem Untergrund in beiden Zeilen weisen auf identische Nukleotide hin. Alle anderen Farbkombinationen stellen unterschiedliche Nukleotide dar. Das Alignment wurde mit ClustalW und Boxshade erzeugt. Die Sequenzidentität liegt bei 91,3 % (ClustalW).

C. pubescens15502	1	MATSSYCATNGVELFNHYGNSRLIRFSSINTSSKSLFTRTSNFRRAKRCFYKSTLSE
S. tuberosum	1	-----MATANGAHLFNHYSSNSRFIHFTSRNTSSKLFLLTKTSHFRPKRCFHVNNLTSL-
C. pubescens15502	61	PKQKVQDPITDLGAESDLSSFPADDAASTASSIKYHAEFTPLYSERFELPKAFFATAQSV
S. tuberosum	54	--BKIHHPITEQGGESDLSSFPADDAASTSSIKYHAEFTPVFSPERFELPKAFFATAQSV
C. pubescens15502	121	RDSLLINWNATYDIYEKLNMQAYYLSMEFLQGRALLNAIGNLELTGAFAEALKNLGHNLE
S. tuberosum	112	RDSLLINWNATYDIYEKLNMQAYYLSMEFLQGRALLNAIGNLELTGAFAEALKNLGHNLE
C. pubescens15502	181	ENVASQEPDAALGNGGLGRLASCFLDSLATLNYPAGYGLRYKYGLFKQRITKDGQEEIA
S. tuberosum	172	ENVASQEPDAALGNGGLGRLASCFLDSLATLNYPAGYGLRYKYGLFKQRITKDGQEEVA
C. pubescens15502	241	EDWLEIGSPWEVVRNDVSYPIKFYGVSTGSDGKRYWIGGEDIKAVAYDVPIPGYKNTTT
S. tuberosum	232	EDWLEIGSPWEVVRNDVSYPIKFYGVSTGSDGKRYWIGGEDIKAVAYDVPIPGYKTRTT
C. pubescens15502	301	ISLRLWSTQVPSADFDLSAFNAGEHTKACEAQANAEEKICYILYPGESEEGKILRLKQQY
S. tuberosum	292	ISLRLWSTQVPSADFDLSAFNAGEHTKACEAQANAEEKICYILYPGESEEGKILRLKQQY
C. pubescens15502	361	TLCASLQDIISRFEERRSGDRIKWEFPEKVAVQMNDHTPTLCIPELMRILIDLKGLNWN
S. tuberosum	352	TLCASLQDIISRFEERRSGDRIKWEFPEKVAVQMNDHTPTLCIPELMRILIDLKGLNWN
C. pubescens15502	421	EAWNITQRTVAYTNHTVLP EAEKWSYELMQKLLPRHVEIEAIDEELVHEIVLYKYGMD
S. tuberosum	412	EAWNITQRTVAYTNHTVLP EAEKWSYELMQKLLPRHVEIEAIDEELVHEIVLYKYGMD
C. pubescens15502	481	LEKLEELTMMRILENFDLPSSVAELFIKPEISVDDDTETVEVHDKVEASDKVVTNDEDD
S. tuberosum	472	LEKLEELTMMRILENFDLPSSVAELFIKPEISVDDDTETVEVHDKVEASDKVVTNDEDD
C. pubescens15502	541	TGKKTSAKTEAAAEKDIKKTTPVSEPAVILPKKVRMANLCVVGHHAVNGVAEIHSEIVK
S. tuberosum	532	TGKKTSAKTEAAAEKDIKKTTPVSEPAVILPKKVRMANLCVVGHHAVNGVAEIHSEIVK
C. pubescens15502	601	EEVFRDFYELWPEKFQNKTNQVTPRRWIHFCNPALSSIIITKWTGTEDWVLKTEKLAELQK
S. tuberosum	592	EEVFNDFYELWPEKFQNKTNQVTPRRWIHFCNPPLSAIIITKWTGTEDWVLKTEKLAELQK
C. pubescens15502	661	FADNEVLQTEWREAKRSNKIKVASFLKEKTGYSVVPDAMFDIQVKRIHEYKRQLLNIFGI
S. tuberosum	652	FADNEVLQNEWREAKRSNKIKVVSFLKEKTGYSVVPDAMFDIQVKRIHEYKRQLLNIFGI
C. pubescens15502	721	VYRYKKMKEMTAERKSNFVPRVCIFGGKAFATYVQAKRIVKFITDVGATINHDP EIGDLLKVVFPDYNVSVAEELLIPASDLSQHISTAGMEASGTSNMKFAMNGCIQIGTLDGANVEI
S. tuberosum	712	VYRYKKMKEMTAERKTNFVPRVCIFGGKAFATYVQAKRIVKFITDVGATINHDP EIGDLLKVVFPDYNVSVAEELLIPASDLSQHISTAGMEASGTSNMKFAMNGCIQIGTLDGANVEI
C. pubescens15502	781	LKVVFPDYNVSVAEELLIPASDLSQHISTAGMEASGTSNMKFAMNGCIQIGTLDGANVEI



<i>S. tuberosum</i>	772	LKVVFVPDYNVSVAEELLIPASDLSEHISTAGMEASGTSNMKFAMNGCIQIGTLDGANVEI
<i>C. pubescens</i> 15502	841	REEVGEENFFLFGAQAHEIAGLRKERADGKFVPDERFEEVKEFVRSVFGSYNYDELIGS
<i>S. tuberosum</i>	832	REEVGEENFFLFGAQAHEIAGLRKERADGKFVPDERFEEVKEFVRSVFGSYNYDELIGS
<i>C. pubescens</i> 15502	901	LEGNEGFGRADYFLVGMDFPSYIECQEKVDEAYRDQKRWTKMSILNTAGSYKFSSDRTIH
<i>S. tuberosum</i>	892	LEGNEGFGRADYFLVGMDFPSYIECQEKVDEAYRDQKRWTKMSILNTAGSYKFSSDRTIH
<i>C. pubescens</i> 15502	961	EYAKDIWNIEPVKTP
<i>S. tuberosum</i>	952	EYAKDIWNIEAVETA

## A 15

Potentielle Nukleotidsequenz des Kandidatengens 8248 aus der *C. pubescens*-Akzession PI 235047

&gt;Kandidat 8248

ATCTGAACATTATCGTGCCTCATCTTTATAATTACAACAATATATTTTTGCGTAATTTTTCGCAAGTGAATTTTGAGAAAGGTG  
 GTGTATATATAGTCTTATCTCTATCTTGTGACGACATAGACTTGATTCAATTTCAAGAAAATTAATAATTCATAGTTGTGACACT  
 CATTGTGCTAGCTAGATCAAATGGCCAAAGTTATGAATGGTGACCAAATGTGTTGGATGGAAGTTGCTGTTCCAGCAATTCTCATGA  
 CACCAAATAAGATTCTGTTTCTCCTCAGCTTGAAGTATATTTGAAGAAGAGGATTTTGAGGACTCTGATTTTCAACAAGAAGGG  
 AATTCTATTTTTTCCCTATAGATTTTAGTTGAATTAATTAGATGTGAATTTTTCTTTTTCATCATTTGAGATTTGTTAAGAATTA  
 AGATCTCTATGGTTCTGTTTTTGTGTTATGTAACATTTTTTCGATTCCAAAGAATTCATTTGATTAGTTG

Ein Sequenzcontig des *C. annuum* cv. CM334 Genoms, der einen homologen Bereich zu dem Kandidaten 8248 aufweist.

Der homologe Bereich im *C. annuum* cv. CM334 Genom ist grau hinterlegt. Punkte “...” am 5’ und 3’ Ende des Contigs deuten an, dass hier nicht die komplette Sequenz dargestellt ist sondern nur Teil. Gelb markiert ist das identifizierte Effektorbindeelement des Kandidaten *EBE<sub>AvrBs4</sub>Cand8248*.

...GTGCCATTTTACGTGTCGTTATTTTCTTTTAGTCCGTCCTAAAAAGAAATGTCATCTTTCTTTATTTGATAACTATTTAATAACA  
 ACATCCTCATTTTATTTTACTGGATCCCCTTAATAAGAAAAGACAACATAAAAAATAACTTTAAGAAGAACATTTTATAAACT  
 TTACAAAAATCATTACTCATTTTTTAACTCCGTATTCGATCAAAAAGGTGATACATAAAAAATAAGAACGGAGGAGTATAATAT  
 ATGAAAAATCTCTTTGTGATTTGTCAGATTCATTATTGTATGTTAAATTTGTGGTCTTAAAGTTTTTGTAGTGATTTTAAAGCGT  
 GTGCAGATTTCCCATGTTTACATGCTTTTACTTTAAAAATTTGGCAATATTGAAAAATGACAATATTCAGTGATGTTTCCACTGTC  
 TTTTACCCTCCACCCCTCACCCGGTGGCGAAGCTATCTTTGCTTCAAAAATTTTGAACCTTCCTTCGATAAAAAATTTATA  
 TTATTTTATATGATATAAAATACCTTTTACGTATATATAGTAAATGCAAAACCACGTTTCGACTAGTTTGTATATTACTTCTAAACT  
 CAATGAAATTTCTGACTCCGCCACTACCCAACCCCAACCCCTTCAATTCTTCATTTCTCTCACTGCTCCACCTATATAATATTAG  
 TCCCTCCCTCCCTTCCCAACACCACAAAAATACACTCAAAATATTAGTCCCATGATTGATTATCTGAACCTTTTATCTTGTCTCGG  
 TGAAAGTTGATCTTGCAATTACAACAACATTTCTCTCACCTCCCTCCACCTTCAAATATTATCCCTCCCTTTTCTATCTCTCAT  
 ACATTTCAAGAAATATACTCAAAATAAAGTAGTAGTAAAGACTCTTTGGCTAAAAAAATCTTAATAGTTTCGGTTGGTTGGCTTTC  
 TACTATCTGAACCTTTTAACTTGTGGTGAAAGGTTCAATCTCTTAAAGCTAAGTTTCATTTCTGCAATTACAATAACAATCTCAAA  
 CACCACAAGAAACACACTCAAAAGTAGTACCAAGACACTCAGAAAATCTTAGTAGTTTCATCTGGTTGGTTATCTGAACCTATTATC  
 GTGTTTCTATTCTTATAATTACAACAATATATTCTGCGTAATTTTCGCAAGTGAACCTTTGAGAAAGGTGGTGTGTATACCGTCTTA  
 TCTCTATTTGTGACGACATAGACTTGATTTAATTTGTGACGACATAGACTTGATTTAATTTCAAAAAATTAATAATTTTCATAG  
 TTGTGACACTCGTTTGTAGCTAGATCAAATGGCCAAAGTTATGAATGGTGACCAAATGTGTTGGATGGAAGTTGCTGTTCCAGCA  
 ATTCTCATGACACCAATAGATTCTGTTTCTCCCAAGCTTGAAGTATATTTGAAGAAGAGGATTTTGAGGACTCTGATTTTCA  
 AGAAGAAGGAATTTCTATTTTCTCTATAGATTTTAGTTGAATTAATTAGATGTGATTTTTCTTTTTCATCATTTGAGTTTGTG  
 TTAAGAATCTCAAGCTCTATGGTTCTGGTTTGTGTTATGTAACATTTTTCGATTCCTCAAGAAATTAATTGATTAATTGATGAA  
 AGTGTGCAATTTATGTAAGGCAAAATGATTTTCTAGACCCCTGTACTATGTCGGTTGGTTTGTAGTTGGACATTTCTACTTACA  
 TGTTTGTCTCTGAACACTTAAACCATTAATAATGCGATATTTGTCATCTTTGACCATTGACCGAGCCTATGTGGCATTCAAATG  
 CTGACTAGGATAAAGAGAGTGTAGTCACACGCCTGAGGGTGCAATTTTGTATCAATTTTATATTAATAATATTTTT  
 TAAAAAAAAGTGGCACTCAATTTTTTTCAGTTTTTAAATTTTAAATATTTTAAATAATTAATAATAATAA...

## A 16

Sequenzen der *Bs4C*-Homologe im Scaffold 946 auf Chromosom 1 im assemblierten Genom von *C. annuum* cv. CM334

Dabei sind das Startkodon grün, das Stoppkodon rot markiert und die potentiellen *EBE* gelb hinterlegt. Dick schwarz markiert sind vorzeitige Stoppkodons und in blau das zweite ATG in den Sequenzen von *CaBs4C.3* und *CaBs4C.4*.

&gt;CaBs4C.1 (Contig 011445 – Teil des contig 184452, Scaffold946)

CCGACTATAGGAGAGTCAAGTCCGAAATGTTTTATCAAATGGGGCTACGAAAAACATTATTCCTAATTACACTAACGTCCCAGT  
 GGAAAGTACCTTCATGCAGTAGTATGGTACAATATTTAAAGCGATGAACTTTTAAATACCAAAAAATAGCCCTCCAGCCTTC  
 ACCAATTAGGAATATAAAATATTCTAAATCAACACTACTGGATAAAATGGATATATTGACCTAATATATTTAACTATATTTATA  
 CTCTTTCTCTAAGCTTCGAGGAAAGATCGCGTACGTGATCTTAAATGACTGTCTATAAATAAAATTAATAATGACACACACAG  
 CATACTATACCAACCATATAATTTAGGAAAAGATTCAACAGAATATAATATGGAGTTTGATCTAAGATACTTCATCTTAATTTT  
 GGCTAACATGCTCAATCAATATTATCCATTTCTAATAACTGGGATCCTTTCCATATAATTGGTAACCCCATTCAGTTTTCATGA

CATTATGAATAGGCTCTTCTTTCTTTTCATATTTTCTTCCATCTTCTCCATCAGTATGACACTTCATCATCCAATTATACGA  
ATACGTGTAAAGTATTACTTTCTCCAAGTCGTTTAAATATCTTGTGCTAGCTTCTCTTCTACTCCCCAAATGTTGTTCTGGTACTT  
TTCTTCAATTATCATTACATTTCTTCTTATGTTCTTCTGGATTGGCAATACATATGGGCTAGTTTTCGACAAACGGATTTCGCATATTT  
TCTCAACAATAAATATTTCCAGCAGTTAGCATTTTTCATCAATGTGGAGTTTGATAGAAATAATGTCTATCAGCCAGCAGCATGAAACT  
CAACAAATCACCTGCTCTCTCTCGTC**TAGA**ATCTAGATGAATTCATCGATCAAGTATTTTATGTTTCATCCAACGCTTACATTTT  
ATATATGTGTTGTTTCCCTATGAAAATCTTGAAAATAAGGTACGTTTACGAACATAAAGTAAAGATGTCATGATAGTGTGA  
AATTTATGGGTGTCTATTACTACTGTTTATACTAATATATAGCTGTTTAGCGATGTGTATGCATAAATGGATACATCAGCTTTGT  
CAGGTTGATAGATATTCATCATCTGCAGAAGCATTTTCCAGTTAAACCATATTTGTGCTGAATGCTGTGATGCTGTTTAAATTAGTCGG  
CACTGTATGTCGAGTAAAGTTTTCAATAATATTTCTGTTGTTATATAGTATTCTAAGTTTGAACATAATTTTCTTCTTTTCT  
TCTTCTTCTGTTTATTTTAAAGGATATGTGAATTCAGTCAGTACTTGTACCGTTCAACAGTAAAGGTTAAGTTGTTAGGA  
CTCTACAAATATGTCAGTGGGTATGTGACCCCC

>**CaBs4C.2** (Contig 082548 revcomp, Scaffold946)

AACTTCGATCCTTCCATCTTCAGTGATCGTGATGCTCCCAAGCTTTGATACCAGTTGTTAGGATCGATTAGCAGTTCACGCACAC  
TAAATCTGTGGAGAGGATGAAGAAAATTTGGAGAGAGAGTTCTTTAGGAGAGAGAGAGAGAATTAATTTTCGTGGTAAACCCCGTG  
GGTAACCTCCAGGGAGACAGGCTATTTATATTAATGATTTAGTGGCTAACTTTGCGATGGCACAACCTTCCAAATGGCATAAAA  
AGATTGAACGGTTAAGACTTAATACTAACTAAGATTTATAGTCAAACTATATACCATTATTTTGAAGTCAAGTTCATGATGTCT  
CCTCAAGAGGAAATGAAAAAATGGAGAGGAGCTCAATCCCTTTCGGTATATACCGCTGGGTCTGAGTGTGACACTTTTATTT  
TGACCACCTCACCAACCAACAGTGATATTTATGTATGTAATTTGTGGACGATCTTTCCGATGCATGCAGTTGTCTTTTAGTGGACAA  
ACTTTCCAAATGGCATAAAAATACATATATAGTTTCTATGTAGTGGGCTAACTTTGCGATGGCATAAGACTTAGAAGTAGTAAC  
ATACATATATAGTTTCTATGTAGTGGGCTAACTTTGCGATGGCATAAGACTTAGAAGTAGTAACATACATACATGAGGCTGG  
TGTTTTTGGTTCAGCAATGAAAAAGTACCTTCATGCAGCAATAGGGTACAATATTATTAAGCGATGAAACAATTAATA**TGCCAAA**  
**AATAGCCCTCCA**TGCTAGTAGACAGCCATGACAACATATGTAAAGCCGTTCAATCTCGAGAAAAATTTGGGCCAACCTATTTATA  
CACGCTCAACTTCGAGAAAGAGCAGAATGTTAGAATAATAATTTACGCGTAGATATAAATAATGAAAAACACACAGCAGGAA  
GCATTATCACTTGTATATATCCACTCAAAATTTTCAAAAAGCTCAAGAAGAA**ATGG**AGTTTGATCTTAGATCTTGATCTTAAT  
TATAGTTAACTTCTCAATCCATATCACCTTCTGAAAATTTGGGATCCTTTCCATATATTTGGCAATCCCATTCCAAGTTTCATCA  
CATTTATGAATAGGCTCGCCTTTCTTTTCTTATTTTCT**CTTCATTTTCTCCATCAGTACAACTTCATCATCCAATTATACGA**  
**ATACGTATAAGTATCTTTCTCCAAGTCTTCAATTTCTCCATCAGTACAACTTCATCATCCAATTATACGAATACGTAT**  
**AAGTATTACTTTCTTCAAGTCGTTTAAATCTTGTGCTAGCTTCTTCTTACTCCCAAAATGTTGTTCTTGGTATCTTTTCTTCA**  
**TCACCATTACACTTTCCTCATGTTCTTCTTGGATTGGCAATGCATGGGCTAGTTTTCGACAACCGACTCTGCATATTTCTCTACA**  
**ATATTTCCAGCAGTTAGCATTTTTCATCAATGTGGAGTTTGACAGAAATACTGTGAGCCAGCAGCATGAAACTCAACAACCTCGCCT**  
**GCTCTCTCTCGT****TAG**ACGAATTCATCGATCAAGGATATATTTCTTGTTCATCAACGCTCTGTATTTTCATATATGTGTGTTGTTT  
CTGTATTGAAATCTTGAAAAAAGGTACGTTTACGTATATAACAGGTAAAGATGTCATGACAGTGTGAAATTTATGGGTGCTGT  
TTACATGGACTACTCTTTTATACCAATATATATATAGCTGTTTAGCCATGTGTATACATAAATGGATACATGAGCTTTGTACCTT  
GATGGATATTCATCATCTGCAGAAGCATTTTCCATGTTAACTATTTGTGCTGAATGCTGCTGATACGTTTATTAGTCACTTGTAT  
GTTTGAGTAAAGCTTTTCTATAATATCTTCTTGTATTGCAAGTATCTAAGTTTATTTATTTTGAAGATATGTAAATTT  
TGGTCAGTACTTGTGCTGATCAACAGTAAAGGCTAAGTTGTTCCGATCTGTAATAATGTTTTGAGTGTGTGCGAATCTTCA  
AAAATAATGTATTTTGAAGAACCAACAGGGTGGGCAATATTTGTGAAGAGTCTGAGCAACTAG

>**CaBs4C.3** (Contig 214709, Scaffold946)

TTACACAACCTGTCTGGTGTTTTATTTGGTTCAGCAACGAAAAAGTACCTTCATGCAGCAATATGGTACAACATTAAAGCGATGAAA  
CAATTAAAA**TGCCAAAATAGCCCTACAT**GATAGTAAACCAACCATTAACAATATTTGTAAGAAATTCATCTAGAGAAAAATTTGGA  
GCCAATCATTTTATACACGCTCAACTTCGCGAGGAAAGAGCATAATGTTGAAATATAAGTTTAGGCGTATGTATAAATAATGAAA  
ACACATTGCTATATTTCTACTCAAACATACATA**ATGG**AGTTTGATCTCTGATATAGT**TAAC**ATACATTTCTCAATCCATATCACCTT  
CTGATCAAAATTTGATCCTTTCCATATATTTGGCAATCCCATTTCCAAGTT**TCATG**ACATTATGAATAGGCTCTCTTCTTCTCTC  
TTATTTTCTCTTCAATTTTCTCCATCAGTATACACTTCATCATCAATTTATACGAATACATATAAGTACCCTACTTCTTCTCAGCA  
TTTCTCCAAGTCTTTAATATCTTGTGCTAGCTTCTTCTTACTCCCAAAATGTTATTTTGGTACATTTTCTTCATCACCATT  
CATTTGACCTCATGTTCTTCTTGGGTGGCAATACATGGGTAGTTTGGACAACCGACTTTGCGTATTTTTTCTACAATATTTCCA  
CGAGTTGATCATTTTCATCAATGTGGAGTTTGATAGAAATAACGTAAGCCAGCAGTCTGAAACTCAAACCTACTCATTTACTCTCTCT  
CGT**TAG**TTGAAGCATCGATCAAGGATATATTTTCATGTTTCAATGTTTCAACGCTTTACATCTCATATATATGTTGTTGTTCTGT  
TGAAATCTTGAAATAAGGTATGGTTACCTACTATAACAGGTAAAGATGACATGATAGTGTGAAATTTATGAGTGTCTATTATA  
TGGAGTACTACTCTTTTAGTAATATATAGCTGTTTAGCAATGTATATACAGAAATGGATACATGAGCTTCGTCACGTTGATAGATA  
TTCATCATTTGGAGAAGCACTTTCCATGTTAACTAGTGCTTTGCATGCTGATGCTGTT

>**CaBs4C.4** (Contig197639, Scaffold946)

TATTTTTTGGGTCAAAAATATTTTCTAAAATGTATTTTGGCGCTTTACCTAAACACTAAAATGTATTTGTTGAAAAATATTTT  
TCTCTTTCACCAATCAAAATACTAGAAAATATTTTCAAAAAATATTTTCTATCCACTAACCAAAATATAAGAAAAACAGTAAAAAAT  
CAATTTGTTTTTCAAGGAAATCTTTCTTCTGACCAACACACCTTAGATTAACAAAGAAAAATCTTTAAACATTTTCCGCAAC  
TTGAGAAGTACATATTTAGAGCAGGAACCCCAACCGCTACCATAAAGCTTATAATTTATAGAAAAATTTACCACAAATTTAAAT  
ATTAACAGGTGAAAATTCATAACAAAATAGTAAAATGCTTTGAGATAGGAGAGATGCTACAGGTTGTGAGTTTAAACCTTGTAG  
AAATTTTTTTTTTCTCCGTTTAAATTTTTTTGAGCTTTATTTATTTCACTGGAATCCACATGAAATTTGGATCCGCTTTGTATTGA  
ACTCTGGAAAGCTAGAAATAACATGGTTCAAACCTCAAGAAGATGCTAGTATGTTAACGTTCAAGTGGACAACTTTCCGATGGCA  
TACAAGTAATTATATACTTGTATGGTTAGTGGACAACTTGGTGTGTTGGTTTCAGCAATGAAAGAGTACCTTTCATGCAGCAATAG  
GGTACAATATTTGTTCAAGCGCTGAACAATTAATA**TGCCAAAATAGCCCTCCAT**GCTATCAGATCACCAGGCATGAAGCAATTC  
TCTAGAGTAATTGAATTCATTAAGAATACAGTTTAAAGCGATGTATAAATAAATAATGAAAAAGCTCAAAATTTTCAAAATATA  
TATATATAGCATTTTCAAAAAGCTCAAGAAGAA**ATGG**AGTTT**TGAT**CTTACATACCTTTATCTTGATTATAGTTAAACATACATCTC  
AAATCCATGTCATCTTCAGAAAATTTGGGATCCTTTCCATATATTTTGGCAGTCCCATTCAAGTAT**ATG**ACATTTATGAACAGGCT  
CGTCTTCTTTTTCATATTTTCTTCAATTTTCTCCATCAGTATGACACTTCATCATCCAATTATACAAATACGTATAAGTATTA  
CTTTCTCCAACTCTTTAATATCTTGTGCTAGCTTCTCTTCTACTCCCAAAATGTTGTTCTGGTACATTTTCTTCATTTATCAT  
ACATTTTCCCCATGTTCTTCTTGGATTGGCAATACATGGGCTAGTTTTCGACAATGGACTTTGCGTATTTTCTCTACAATAATATT  
TCCAGATTTGCTATCTTTCAGCAATTTGTCACGTTGATAGATATTTATCATCTGTCAGAAACATTTTCCATGTTAACTATTATGCTGA  
AGGCTACTGATGCTTTT**TAAT**GAGTCACTTGTACGTTTGACTAAAGCTTTTCTACAATGTTTTTGTGTTGTTAGTTTCTTCTGCG  
GGTACTAAAATCTGTATTAATACACAACAAATTTTCAACAAGTTCATGTATGAACAAGAACAGCTTTGAAGTTCTTTTAAACCTT  
TCAACACAAGATTATTACTAATTTATCTAGAAAGTGTGTAGATATCAGAAAATAGATGAAACAGAAATTTTATGATCAAACTCTCC  
GAGTTCAGGATGCTCTTAAAGGAATAATCCCTCACGATACGAGGTTATGGAATTTTCTCCCAAGTAATAAGTGGCCTTCAA  
TCCAAATATAGGGTATCTCAAAATATTGGATTCCACGAACGCACTCAACAATTTGAATGATCACACAGAATATTT

## &gt;CaBs4C.5 (Scaffold946)

TATCGAATCGATCGATACCCGATATGTACACCTATT TTAAGGTGAGCTTTTCCCAACCAACAAAAATATCTTGGTTGTCCCTA  
TTATGTTGGGTCTTTGTTGGCATGCCAAACACGTGGATGAGCAAAGCAATTTACAAAACATTATAGTAAATGATAATTTTGCCCT  
TTACTTGGTTGACTATCTTGTCTTTGATTGTTGCTATCTATATATATTCATTTAGAAAACAAAACAAAATATACTCAAACATAAAG  
ATACAAAAGTTTACCAAGAATGGAGTTTGTATCTAAGGTACACAATATTAATTATAGTTAACATTCTCAAATCAATATTTATTTCAA  
ATGAATGGGATCCCTTTTCATATATTTGTCCAACTCACTTTCATGGCATTATGAATAGGCTTTTCTCTCTCTATTTGTTTTCTTTC  
ATTTTCTCCATCACTAGTATAACAATTCATTATCCAATGAGAGTACATATAAGTTCAACATATTTCTCAAAGTCCCTTAATATATT  
ATTTTAACTTCACTTCTACTTCCACAAGTGTTATTTTGGTACATTTTCTCATTACTATTATGACATTTCTCTCAATTGGATT  
CCAATTTTAAGGAATGGTCTTTATATATTATTTCAATAATTCAGGATTTAGTATTTTCATCAATGCTGAGCTTGATAGAAATAAT  
GTGAATAACCAGCAGCCCCAAGCTCAAACAATTCACCTACTCTCTCTGTCTAAATTAAGCTATGTTTAAAGATAAGACAGAGGCGG  
ATTCAAAATTTTAAAGATGCAGTTTATTGGATTATGAAGAAGTTATTGTACCTAT

## &gt;CaBs4C.6 (Scaffold946)

TATGTTTCAAATATGTGTTGTTGTTGTTTCTTTTATACAGTACATTTTCTGTTGATTCTAGTGATAAAATATGCCTGCGGAATT  
TTCAGTCAGATCTTAAAATCCAGCTTAATCATGAAGCATTGAATAAAAGAGTTAAAGTTAGGAAAAAGAGCATATGAGAAAATAG  
ATAGAGTGCTGAGATATTTGGTGATCAAGTTAAAGCCTTCTTTGAAAATTAAGGCAATTATTTTGAGAATCGTGAGAATAACTTGG  
TCAGCATTTGAAAGACATGCTCAATTAGGTCAAATAGTGGATGTGTGATCCTACATCAAAAGGAATAGAAAAATTAGCTCCACGA  
AAGCATGAGGAATGTAACCAAGAGATGGAGTTTTCAGAGGAAGTAGTGACGCACAGTCTCAGTTAAATTTAGTGCTTAAAAAT  
CACAAGTGATTTTCATCTCTCACTTTTATATTGCAGATATGTACTTGCATTTTCAAGGATTGATATGACGAAGGCATTTTCATTCTA  
CTATCCAAATATTGTGTCATCCTTTGTTTACCCCATTTGAGCTTTAGTCCATTTTATTTTCATACCCCTTTATTGGAATCAGATA  
GAGTCGGTAAAGCTCAAAAGAAAAGTGTGAGCAAAACAATAGAGTCAGAAAGTTTCAAAAGAAACGAGACAAAAAGTGTCTAA  
AAGAGAAAAAGAAAAGCGTCGAGAAAAATAAAGTCAAAAAGGTCCCAAAGATAAAAAGAGTCAACAAAAGAAAAAATCAGAATC  
AAGCAAAAGAACAGCTGCCAAAATATGCGTGACCTGGTCTCTCTCTTTTGGGGGTGAGATACGTAGGCTGCCCTACCCGGGGCT  
CAGTCCAATCAAATAAAAGTTTAGAATTTCCAGTCAAAAGAACTGGGGCATGAGTTAATCCTCATTGTTGAGTCGATTCCAAA  
AGTTGTAAGTCCTA

## &gt;CaBs4C.7 (Scaffold946)

GGTTAGGTGACAGGTTGGTCTCTGGTCCCTATTGGATTGAGATAACCATCATGGACAGGCCTTGGTTTGGGTCATGACAAAGTTT  
GTATCAAATCCAGCTTTGGTGTCATTGAATGTCCCAAAGCTATTTCAAGTAGAGTCTCTTTTATGGGTATGTAGCGTTCCACACT  
TATAATGGAGAGGTATAAGGAATTTAGAAAATTTTCTTTCTTGTGTTCTATTTTGGCTATAGAGTCTAAGATTGTAAAAT  
CTTCTCTAATCTCATATGTTTGGCTTTTCAGATCATGCCACCTAGAAGATCTGATATTCTGTAGAACTCTTTTGTCTGTCTGAGT  
ACAAATGTGGATGGAGCAGTCTCTACTCATGGAGTTTAGAACTGGCTAGGGTCGCTGCTTCTGATTGCCCTTGTGGAGTTCCACTG  
GTTTCTCTAGCCCTCTTTCAGGCACCTTAAGTTGGCTTATGTAATACTGAGTTTTCAGAGTTGATCCATTGCTTACCCAGTTGGT  
GACTTCTCAAGCCGAGCATGCTAATTTTGTGCTCCATCTTCTATGGTTAATAAGGTTGGCCAATTTATAAGGTTGAGTTCTCCAA  
CTTTCCTGTTCTTAAAGTCGAGGAATATCTCAAATTTTCATTGATGAGATAGAAAAAAAATTTGCTATAATGCATGCTATTGA  
TTTGTAGGGTGTGGAGTTTGTGCTATCAGTTGAATTATGTGACGTATCAGTGGTATTATAGGGACGATGATGACAAGTCCACTC  
TATGGGATGACTTCTAGTGGTTTCTTGAATATTTTCTTTCAGGAGTTGAGGGGTGAAGGCTAAGGAATTTGAGAATTTAAAG  
TAAAGCAGGATGACAGTAAAGAGTACGCCTTGAAGTTTCATCAGTTGTCGTGTTATGCTATAGAGTTAATGTCTATCATGAGGAC  
TAAACTAAGAAAATTTTCTATAGGGTTGTCCCGTAGTTGATTTTGGAGAGCAAGGCTGCCTTGTGGAACAGGATATGGACATCT  
CTATACTTATGGTGTATATCCAACAAGTGAAGAGAAAAAGATAAAAGAAAGCAGAGATGAGTAAAGGCAGAATAAGAAGTTTCAA  
TA

Proteinsequenzen der *C. annuum*-Homologen mit durchgehendem Leserahmen.  
Die Duplikation in CaBs4C.2 wurde mit dicken blauen Buchstaben hervorgehoben und in der  
Sequenz CaBs4C.2\* wurde sie deletiert.

## &gt;CaBs4C.1

MEFDLRYFILILANMLKSILSISNNWDPFHIIGNPIPSFMTFMNRLFFLFIFSFIFSITRMTLHHPPIIRIRVSITFSKSFNLCCLAS  
LLLPQMLFWYFFFIITLSSCSWIGNWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVISQQHETQTNHLLSLV

## &gt;CaBs4C.2

MEFDLRYLILIIIVNILKSISPSNWDPFHIFGNPIPSFITFMNRLAFLFLFS**FIFSI**TR**RTL**LHHP**PIIRIRISITFSK**SFISITR  
TLHHPPIIRIRISITFSKSFNLCCLASLLLPQMLFWYFFFIITLSSCSWIGNAWASFRQPTLHIFSTIIFPAVSIFINVEFDRNTV  
SQHETQT TRLLSLV

## &gt;CaBs4C.2\*

MEFDLRYLILIIIVNILKSISPSNWDPFHIFGNPIPSFITFMNRLAFLFLFSFIFSI**TR**TLHHP**PIIRIRISITFSK**SFISITR  
LLLPQMLFWYFFFIITLSSCSWIGNAWASFRQPTLHIFSTIIFPAVSIFINVEFDRNTVSQQHETQT**TR**LLSLV

## &gt;CaBs4C.5

MEFDLRYTILIIIVNILKSIFISNEWDPFHIFVQLTFMAFMNRLFLLLLFSFIFSITSITIHYPMRVHISSTYFSKSFNIFLTSFL  
LPQVLFWYIFFITIMTFPSNWSNFKEWLSYIISIIPGFSIFINAEIDRNNVNNQQPQAQTIHLLLPV

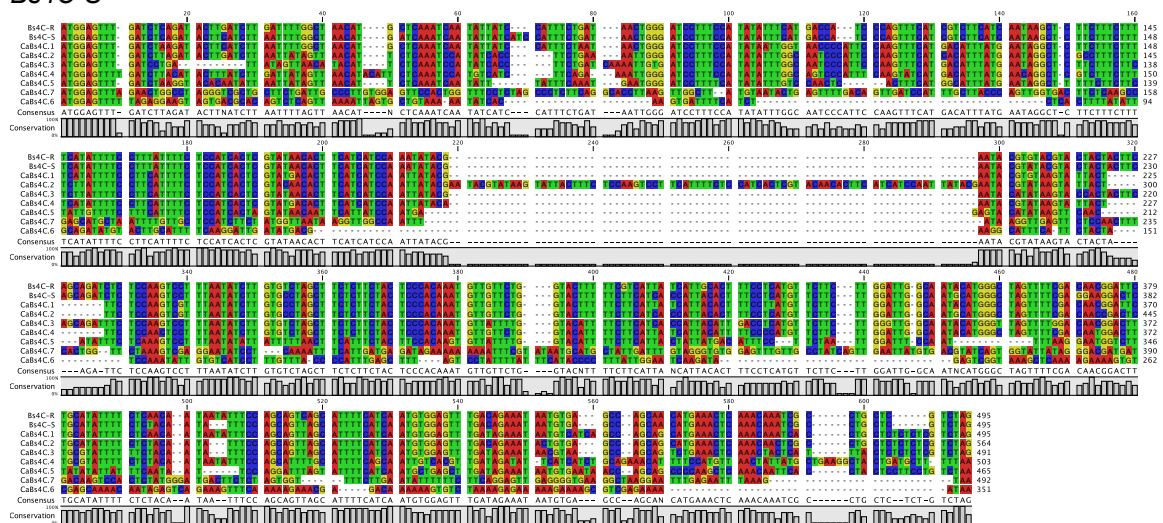
## &gt;Bs4C-R

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRVRTTSADLSKSFNIL  
CLASLLLPQMLFWYFFVIIIALSSCSWIGNWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

## &gt;Bs4C-S

MEFDLRYFILILANMIKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRTTTSADLSKSFNI  
LCLASLLLPQMLFWYFFFIITLSSCSWIGNAWASFRGRTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

### Alignment und der Vergleich Sequenzen *CaBs4C.1-CaBs4C.7* (Start-Stop) mit *Bs4C-R* und *Bs4C-S*



### Vergleich Sequenzen *CaBs4C.1-CaBs4C.7* (Start-Stop) mit *Bs4C-R* und *Bs4C-S*

		1	2	3	4	5	6	7	8	9
Bs4C-R	1		95.58	88.76	73.71	80.28	75.05	67.59	38.48	32.23
Bs4C-S	2	22		86.86	74.74	80.87	73.66	68.18	38.86	31.84
CaBs4C.1	3	57	67		78.88	82.35	81.34	71.00	37.36	32.16
CaBs4C.2	4	153	147	121		73.92	71.06	61.15	31.91	26.84
CaBs4C.3	5	100	97	90	151		75.77	68.50	38.42	30.52
CaBs4C.4	6	130	138	95	169	126		64.98	35.51	31.25
CaBs4C.5	7	164	161	145	223	160	180		35.80	30.74
CaBs4C.7	8	323	321	332	414	327	345	339		26.79
CaBs4C.6	9	347	351	346	428	362	352	347	369	

#### A 17

##### CaBs4C.2 und Bs4C-R-Vergleich

Bs4C-R	1	MEFDRLYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITR
CaBs4C.2	1	MEFDRLYLILIIIVNILKSI-SPSENWDPFHIFGNPIPSFIFEMNRLAFLELFISFIFSITR
Bs4C-R	60	ITLHHFNIRIRVRTTT-----SADLSKSFNILCLASLLLPQML
CaBs4C.2	60	ITLHHFTIIRIRISITFSKSFIFSITRTTLHHPIIRIRISITFSKSFNILCLASLLLPQML
Bs4C-R	98	FWYFFVIIIALSSCSSWIGNTWASFRQRIHLHFSTIIFPAVSIFINVEFDRNNVSQQHET
CaBs4C.2	120	FWYFFFITITLSSCSSWIGNAWASFRQPTLHFST-IFPAVSIFINVEFDRNTVSQQHET
Bs4C-R	158	QTNRL--V
CaBs4C.2	179	QTTRLLSLV

#### A 18

##### CaBs4C.5 und Bs4C-R-Vergleich

Bs4C-R	1	MEFDRLYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRI
CaBs4C.5	1	MEFDRLYTIILIIIVNILKSIF-ISNEWDPFHIFVQL-TFMAFMNRLFLLLFSFIFSITSI
Bs4C-R	61	TLHHFNIRIRVRTTTSDLSKSFNILCLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWA
CaBs4C.5	59	TIHYP-MRVHISSTY---FSKSFNIFLTSFLLPQVLFWYIFFITIMTFP-SNWNIS---
Bs4C-R	121	SFRQRIHLHFSTIIFPAVSIFINVEFDRNNVS-QQHETQTNRL--V
CaBs4C.5	111	-FKWESLYIIS--IIPGFSIFINAELDRNNVNNQOPQAQTIHLLLPV

**A 19**

Sequenzen von den verkürzten *CaBs4C.3* (*CaBs4C.3\**) und *CaBs4C.4* (*CaBs4C.4\**) ab ihrem jeweiligen 2.ATG und die daraus entstehenden Proteine CaBs4C.3 und CaBs4C.4. Grün ist dabei das potentielle Startkodon und rot das potentielle Stoppkodon markiert.

>*CaBs4C.3\** (Contig 214709, *Scaffold946*)

ATGACATTTATGAATAGGCTCTTCTTCTCTTATTTTCCTTCATTTTCTCCATCACTCGTATAACACTTCATCATCCAATTAT  
ACGAATACATATAAGTACCACTACTTCAGCAGATTTCTCCAAGTCCTTTAATATCTTGTGTCTAGCTTCTCTTCTACTCCACAAA  
TGTTATTTTGGTACATTTTCTTCATCACCATTACATTGACCTCATGTTCTTCTTGGGTTGGCAATACATGGGTTAGTTTGGACAA  
CGGACTTTGCGTATTTTCTACAATATTTCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTGATAGAAATAACGTAAGCCAGCA  
GTCTGAAACTCAAACACTACTCATTTACTCTCTCTCGTC**TAG**

>*CaBs4C.3\**

MTFMNRLFFLLFSFIFSIITLHHPIIRIHISTTSADFSKSFNILCLASLLLPQMLFWYIFFITITLTSCSSWVGNTWVSFGQ  
RTLRIFFSTIFPAVSIFINVEFDRNNVSQQSETQTTHLLSLV

>*CaBs4C.4\** (Contig197639, *Scaffold946*)

ATGACATTTATGAACAGGCTCGTCTTCTTTTCATATTTTCCTTCATTTTCTCCATCACTCGTATGACACTTCATCATCCAATTAT  
ACAAATACGTATAAGTATTACTTTCTCCAACCTCTTAATATCTTGTGTCTAGCTTCTCTTCTACTCCACAAAATGTTGTTCTGGT  
ACATTTTCTTCATTATCATTACATTTTCCCCATGTTCTTCTTGGATTGGCAATACATGGGCTAGTTTTCGACAATGGACTTTGCGT  
ATTTTCTCTACAATAATATTTCCAGCATTTCATTTTCAGCAATTGTCACGTTGATAGATATTCATCATCTGCAGAAACATTTTC  
CATGTAACTATTATGCTGAAGGCTACTGATGCTTTT**TAA**

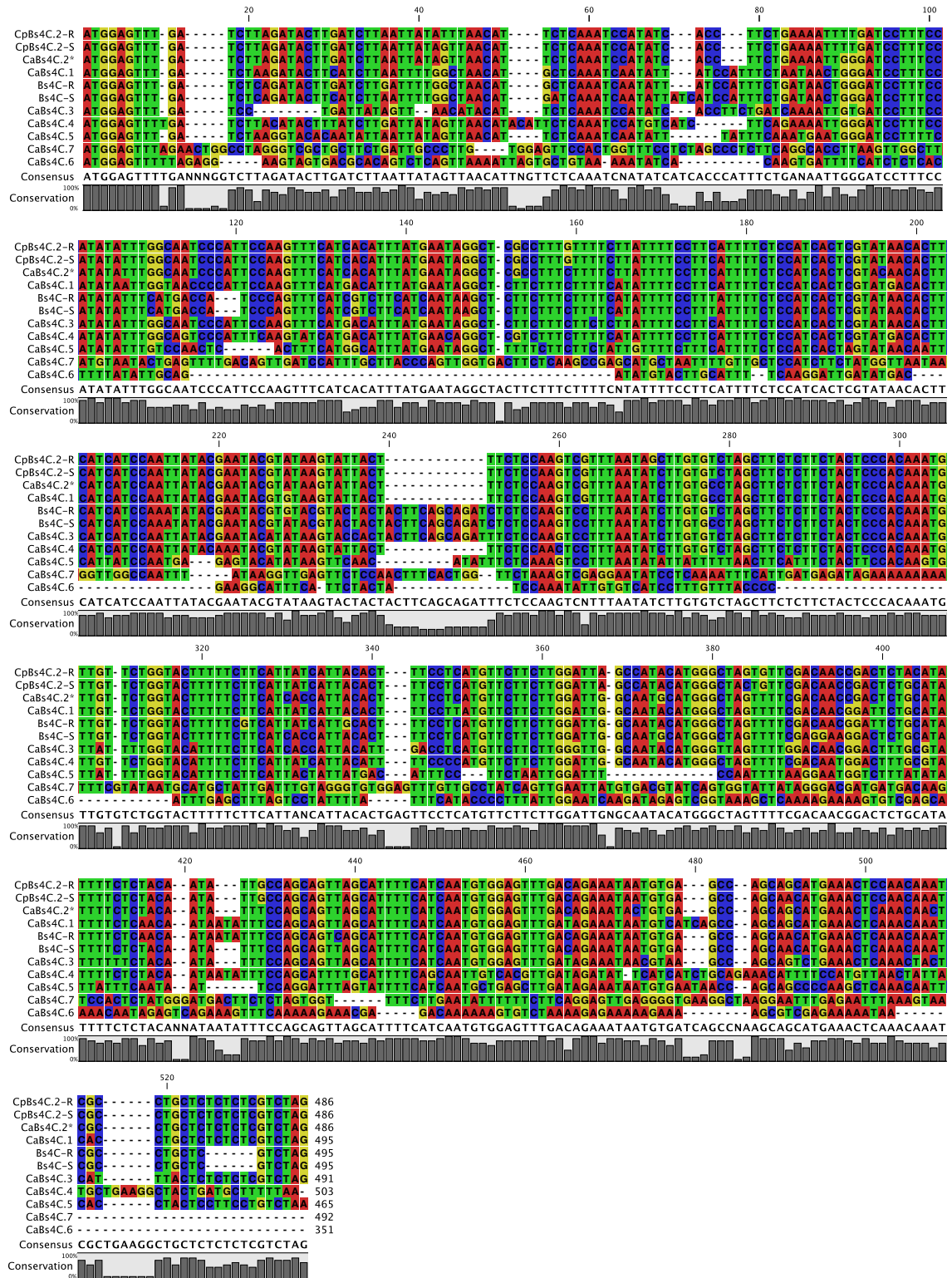
>*CaBs4C.4\**

MTFMNRLVFLFIFSFIFSIITRMTLHHPIIQIRISITFSNSFNILCLASLLLPQMLFWYIFFIIITFSPC  
SSWIGNTWASFRQWTLRIFFSTIIFPAFCIFSNCHVDYSSSAETFSMLTIMLKATDAF-

## A 20

Alignment (A) und Vergleich (B) der Nukleotidsequenzen *CaBs4C.1.*, *CaBs4C.3-CaBs4C.7* mit *Bs4C-R* und *Bs4C-S* sowie einem Deletionskonstrukt von *CaBs4C.2*, bei dem die Duplikation deletiert wurde (*CaBs4C.2\**)

A)



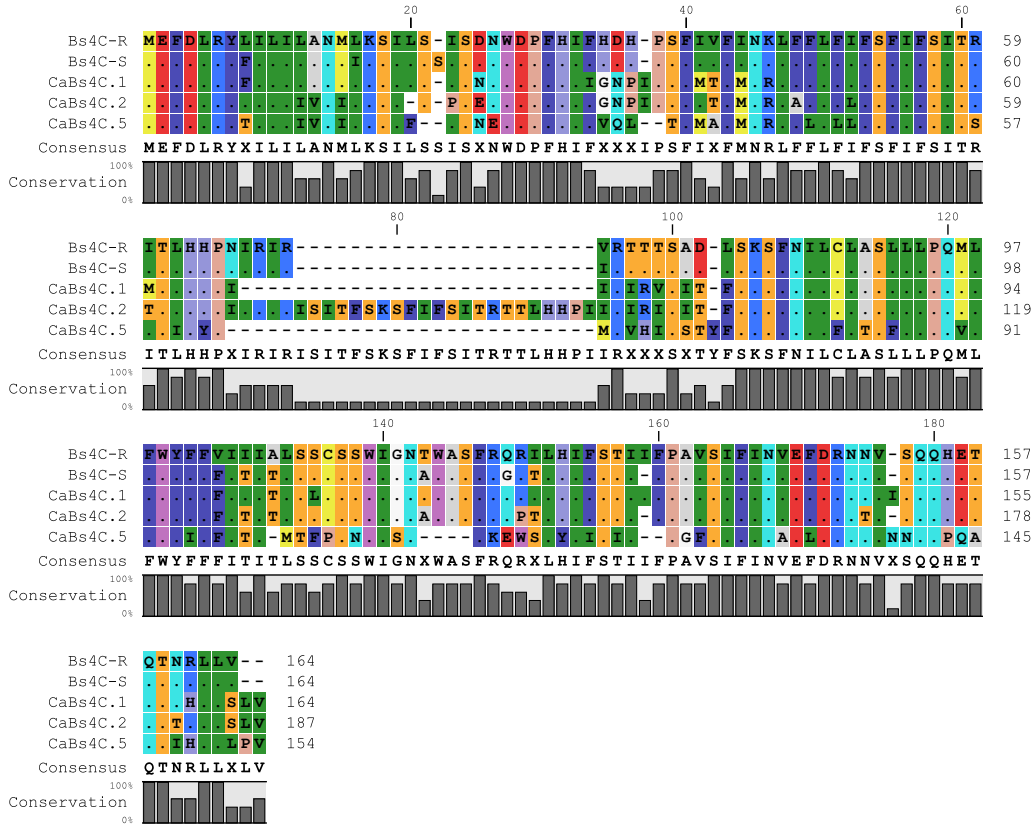
B)

		1	2	3	4	5	6	7	8	9	10	11
CpBs4C.2-R	1		98.97	96.30	90.10	84.33	83.93	84.23	80.12	69.76	36.48	33.60
CpBs4C.2-S	2	5		96.50	90.30	84.92	84.52	84.23	80.51	69.76	36.67	34.19
CaBs4C.2*	3	18	17		91.31	85.12	86.31	85.63	81.30	70.56	35.92	34.00
CaBs4C.1	4	49	48	43		88.76	86.86	82.64	80.23	71.14	37.62	34.77
Bs4C-R	5	79	76	75	57		95.58	80.56	74.00	67.72	38.93	33.86
Bs4C-S	6	81	78	69	67	22		80.95	73.61	68.32	38.74	33.92
CaBs4C.3	7	79	79	72	88	98	96		74.57	68.85	36.23	32.36
CaBs4C.4	8	101	99	95	101	136	138	133		64.65	34.08	32.81
CaBs4C.5	9	150	150	146	144	163	160	157	181		34.91	32.73
CaBs4C.7	10	336	335	339	330	320	321	338	352	343		26.44
CaBs4C.6	11	334	331	332	332	336	337	349	344	337	370	

A 21

Alignment (A) und Vergleich (B) der Proteinsequenzen der *C. annuum* (cv. CM334) Homologe CaBs4C.1, CaBs4C.2, CaBs4C.5 mit Bs4C-R und Bs4C-S mit dem CLC Programm. Alignment zuerst dargestellt, dann Tabelle mit Sequenzidentitäten (oberes rechtes Dreieck/Hälfte) und Anzahl der Abweichungen (unteres linkes Dreieck).

A)



B)

	1	2	3	4	5
Bs4C-R	1	93.33	80.95	69.31	54.17
Bs4C-S	2	11	79.29	70.90	54.76
CaBs4C.1	3	32	35	74.21	59.39
CaBs4C.2	4	58	55	49	52.11
CaBs4C.5	5	77	76	67	91

## A 22

Der Vergleich der 5'UTR-Bereiche der *C. annuum*-Homologen *CaBs4C.1-CaBs4C.5* zeigt, dass diese sehr divers sind und nur kurze homologe Bereiche aufweisen. Grün markiert ist das Startkodon und das potentielle *EBE* gelb hinterlegt.

>*CaBs4C.1*

CCGACTATAGGAGAGTCAAGTCCGAAATTGTTTTATCAAATGGGGCCTACGAAAAACATTATTCCTAATTACACTAACGTCCCAGT  
GGAAAAGTACCTTCATGCAGTAGTATGGTACAATATTATTAAAGCGATGAAACTTTTAAATACCAAAAAATAGCCCTGCGAGCCTTC  
ACCAATTAGGAATATAAAATTTATCTAAATCAACACTACTGGATAAAATTGGATATATTGACCTAATATATTAACTATATTATA  
CTCTTTCTCACTAAGCTTCGAGGAAAGATCGCGTACGTGATCTTAAATGACTGTCTATAAAATAAAATTAATGACACAACCAG  
CATACTATACCCAACCATATAATTTTAGGAAAAGATTCAACAGAATATAATATGGAGTTT

>*CaBs4C.2*

CTAACTTTGCGATGGCATAGAACTTAGAAGTAGTAACATACATTACATACATGGGCTGGTGTGTTTTGGTTTCAGCAATGAAAAAGTAC  
CTTCATGCAGCAATAGGGTACAATATTATTAAAGCGATGAAACAATTAAAAATGCCAAAAATAGCCCTCCATGCTAGTAGACCAGCC  
ATGACAACCTATTGTAAAGCGTTCATCTCGAGAAAAATTGGGCCCAACCTATTTATACACGCTCAACTTCGAGGAAAGAGCAGAA  
TGTTAGAAATATAATTTTACGCGTAGATATAATAATGAAAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCCACTC  
AAAAATTTCAAAAAAGCTCAAGAAGAATGGAGTTT

>*CaBs4C.3*

TTACACAACCTTGCTGTTGTTTTATGTTTCAGCAACGAAAAAGTACCTTCATGCAGCAATATGGTACAACATTAAAGCGATGAAA  
CAATTAATAATGCCAAAAATAGCCCTACATGATAGTAACCAACCATAACAAATATTGTAAAGAAATTCATCTAGAGAAAAATTGGA  
GCCAATCATTTTATACACGCTCAACTTCGCGAGGAAAGAGCATAAATGTTGAAATATAAGTTTAGGCGTATGTATAATAATGAAAA  
ACACATTGCTATATTCTACTCAACATACATAATGGAGTTT

>*CaBs4C.4*

AAGCTAGAAATAACATGGTTCAAACCTCAAGAAGATGCTAGTATGTTAACGTTTCAGTGGACAATCTTTCCGATGGCATACAAGTAA  
TTATATACTTGTATGGTTAGTGGACAAACTTGGTGTGTTGGTTTCAGCAATGAAAGAGTACCTTCATGCAGCAATAGGGTACAATA  
TTGTTCAAGCGCTGAAACAATTAAAAATGCCAAAAATAGCCCTCCATGCTATCAGATCACCAGGCATGAAGCAATTCATCTAGAGTA  
ATTGAATTCATCTAAAGATACAGTTTAAAGCGCATGTATAATAATAATGAAAAAGCTCAAATTTTCAAATATATATATATAG  
CATTTTCAAAAAAGCTCAAGAAGAATGGAGTTT

>*CaBs4C.5*

TATCGAATCGATCGAATACCGATATGTACACCTATTTTAAGGTGAGCTTTTCCCAACCAACAAAAATATTCTTGGTTGTCCCTA  
TTATGTTGGGTCTTTGTTGGCATGCAAAACACGTGGATGAGCAAAAGCAATTTACAAAACATTATAGTAAATGATAATTTTGCCT  
TTACTTGGTTGACTATCTTGTCTTTTGTCTATCTATATATATTTCATTAGAAACAAAACAAATATACTCAAACATAAAG  
ATACAAAAGTTTACCAAGAATGGAGTTT

<i>CaBs4C.4</i>	1	AAGCTAGAAATAACATCGTTCAAACTCCAGAAGATGCTAGTATGTTAACGTTTCAGTGA
<i>CaBs4C.5</i>	1	-----TATCGAATCGATCGATACCCGATATG-TACACCTATTTTAA-----GGTGA
<i>CaBs4C.2</i>	1	---CTAACTTTGCGATGG-CATAGAACTTAGAAG-----
<i>CaBs4C.3</i>	1	-----
<i>CaBs4C.1</i>	1	---CCGACTATAGGAGAG-TCAAGTCCGAAATTG-----
<i>CaBs4C.4</i>	61	CAATCTTTCGATGGCATACAAGTAATTATATACTTGTATGGTTAGTGGACAAACTTGGT
<i>CaBs4C.5</i>	47	C---TTTTC---CAACCAACAAAAATATTCTTG-GTTGTCCCT--ATTATGTTGG-
<i>CaBs4C.2</i>	31	---TAGTAACATA-----CATTACATACATGGGC-----TGGT
<i>CaBs4C.3</i>	1	-----TTACACAACCTTGTCT-----TGGT
<i>CaBs4C.1</i>	31	---TTTATCAAATGGCGCTACGAAAACATTATT-----CCTA
<i>CaBs4C.4</i>	121	GTTGT--TGGTTCAGCAATGAAAGAGTACCTTCATGCAGCAATAGGGTACAATATTGTT
<i>CaBs4C.5</i>	96	GTCTT--TG--TTGGCATGCAAAACA--ACGTGGATG-AGCAA-AG---CAATTTTACA-
<i>CaBs4C.2</i>	60	GTTTT--TGGTTCAGCAATGAAAAAGTACCTTCATGCAGCAATAGGGTACAATATTATTA
<i>CaBs4C.3</i>	19	GTTTTATTGGTTCAGCAACGAAAAAGTACCTTCATGCAGCAATATGGTACAACATTA---
<i>CaBs4C.1</i>	68	ATTACACTAACGTCCAGTGGAAAAAGTACCTTCATGCAGTAGTATGGTACAATATTATTA
<i>CaBs4C.4</i>	179	AAGCGCTGAAACAATTAAAAATGCCAAA--TAGCCCTCCATGCTATCA--GATCACCAGC
<i>CaBs4C.5</i>	143	AAACATTATAG---TAAATGATAATT-TTGCCTT-----TACTT--GGT-----
<i>CaBs4C.2</i>	118	AAGCGATGAAACAATTAAAAATGCCAAA--TAGCCCTCCATGCTAGTA--GACCAGCCATG
<i>CaBs4C.3</i>	76	AAGCGATGAAACAATTAAAAATGCCAAA--TAGCCCTACATGATAGTA--AACCAACCATA
<i>CaBs4C.1</i>	128	AAGCGATGAAACTTTTAA--TACCAAAAAATAGCCCTCCAGCCTTCAACCAATTAGGAATA
<i>CaBs4C.4</i>	236	-----ATGAAGCAATTCAATCTAGAGTAA--TTG-----AATTCAATCTAAAG
<i>CaBs4C.5</i>	182	-----TGA--CTATCTTGTCT-----TTT-----GATTTTGTCTATCT
<i>CaBs4C.2</i>	176	-ACAACTATTGTAAAGCGTTCAATCTCGAGAAAAATTG-----GCCCAACCTATTT
<i>CaBs4C.3</i>	133	-ACAAATATTGTAAAGCAATTCAATCTAGAGAAAAATTG-----GAGCCAATCATTTT
<i>CaBs4C.1</i>	187	TAAAAATTATCTAAATCAACACTA-CTGGATAAAAATTGGATATATTGACCTAATATATTT



```

CaBs4C.4 277 AA-----TACAAGCTTT-----AAGCGCATGTATAAATAATA-----ATGAAAAAG
CaBs4C.5 213 A-----TATATATTT-----CATTTAGAAACAA-A-----ACAAAATA-
CaBs4C.2 228 A-----TACACGCTC-----AACTTCG--AGGAAAGAGCAGAATGTTAGAAAT
CaBs4C.3 185 A-----TACACGCTC-----AACTTCGCGAGGAAAGAGCATAATGTT-GAAAT
CaBs4C.1 246 AACTATATTTATACTCTTTCCTCACTAAGCTTCG--AGGAAAGATCG---CGTACGTGAT

CaBs4C.4 317 CTCAAATTTTTTCAAATATATATA-----
CaBs4C.5 244 -----TACTCAAACATAAAGA-----
CaBs4C.2 269 ATAAATTTTACGCGTAGATATAAATAATGAAAA--ACACAACCAGCAGGAAGCATTATACA
CaBs4C.3 227 ATAACTTTTAGGCGTATCTATAAATAATGAAAA--ACACA-----
CaBs4C.1 301 CTTAAATGACTGTCTATAAATAAAATTAAGATGACACAACCAGCA-----TACTATACC

CaBs4C.4 340 ----TATAGCATTTTCAAAAAAGCTC-----AAGAAGAATGGAGTTT
CaBs4C.5 260 ----TACA-----AAAGTTT-----ACCAAGAATGGAGTTT
CaBs4C.2 327 TTGCTATATAATCCATCAAAATTTTCAAAAAAGCTCAAGAAGAATGGAGTTT
CaBs4C.3 264 TTGCTATAT--TCTATCAAA-----CATACATAATGGAGTTT
CaBs4C.1 356 CAACCATATAATTTAGGAAAAGATTCAACAGA---ATATAATATGGAGTTT

```

## A 23

Alignment der *CaBs4C.1*-Sequenzen aus *C. annuum* cv. CM334 (*CaBs4C.1* CM334) und aus *C. annuum* ECW+ECW-30R (*CaBs4C.1* ECW+30R). Grün hinterlegt ist das Startkodon, rot das Stoppkodon und das *EBE*<sub>CaBs4C.1</sub> in gelb. Beide Sequenzen sind identisch bis auf eine Insertion in *C. annuum* ECW+ECW-30R, die zwischen dem *EBE*<sub>CaBs4C.1</sub> und dem Startkodon liegt.

>*CaBs4C.1* (ECW+30R)

```

ATGCAGTAGTATGGTACAATATTATTAAAGCGATGAAACTTTTTAATACCAAAAAATAGCCCTCCCAGCCTTCACCAATTAGGAATA
TAAATATTCTAAATCAACACTACTGGATAAAATTGGATATATTGACCTAATATACACTACTGGATAAAATTGGATATATTGACC
TAATATACACTACTGGATAAAATTGGATATATTGACCTAATATATTAACTATATTTATACTCTTTCCTCACTAAGCTTCGAGGAA
AGATCGCGTACGTGATCTTAAATGACTGTCTATAAATAAAATTAAGTACACACACCAGCATACTATACCAACCATATAATTTT
AGGAAAAGATTCAACAGAAATATAATATGGAGTTTGATCTAAGATACTTCACTTAAATTTTGCGTAAACATGCTCAATCAATATTAT
CCATTTCTAATAACTGGGATCCTTTCCATATAATTGGTAACCCCATTCAGTTTCATGACATTTATGAATAGGCTCTTCTTTCTT
TTCATATTTTCTTCATTTCTCCATCACTCGTATGACACTTCATCATCCAATTATACGAATACGTTAAGTATTACTTTCTCCAA
GTCGTTTAATATCTTGTGCTAGCTTCTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTCTTCAATTATCATTACACTTTCCT
TATGTTCTTCTTGGATTGGCAATACATGGGCTAGTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTCCAGCAGTT
AGCATTTTCATCAATGTGGAGTTTGATAGAAATAATGTCATCAGCCAGCAGCATGAAACTCAAACAAATCACCTGCTCTCTCTCGT
CTAGAATCTAGATGAATTCATCGATCAAGTATTTTATGTTTCATCAACGCTTTACATTTTCATATATGTGTTTGTGTTTCCCTATGAAA
ATCTTGAAAATAAGGTAC

```

```

CaBs4C.1 ECW+30R 1 -----ATGCAGTAGT
CaBs4C.1 CM334 1 GAAAAACATTATTCCCTAATTACACTAACGTCCTCCAGTGGAAAAGTACCTTCATCGCAGTAGT

CaBs4C.1 ECW+30R 11 ATGGTACAATATTATTAAAGCGATGAAACTTTTTAAATACCAAAAAATAGCCCTCCCAGCCT
CaBs4C.1 CM334 61 ATGGTACAATATTATTAAAGCGATGAAACTTTTTAAATACCAAAAAATAGCCCTCCCAGCCT

CaBs4C.1 ECW+30R 71 TCACCAATTAGGAATATAAAATATTCTAAATCAACACTACTGGATAAAATTGGATATAT
CaBs4C.1 CM334 121 TCACCAATTAGGAATATAAAATATTCTAAATCAACACTACTGGATAAAATTGGATATAT

CaBs4C.1 ECW+30R 131 TGACCTAATATACACTACTGGATAAAATTGGATATATTGACCTAATATACACTACTGGAT
CaBs4C.1 CM334 181 TGACCTAATATA-----

CaBs4C.1 ECW+30R 191 AAAATTGGATATATTGACCTAATATATTAACTATATTTATACTCTTTCCTCACTAAGCT
CaBs4C.1 CM334 193 -----TTTAACTATATTTATACTCTTTCCTCACTAAGCT

CaBs4C.1 ECW+30R 251 TCGAGGAAAAGATCGCGTACGTGATCTTAAATGACTGTCTATAAATAAAATTAAGTAC
CaBs4C.1 CM334 227 TCGAGGAAAAGATCGCGTACGTGATCTTAAATGACTGTCTATAAATAAAATTAAGTAC

CaBs4C.1 ECW+30R 311 ACAACCAGCATACTATACCAACCATATAATTTTAGGAAAAGATTCAACAGAATATAATA
CaBs4C.1 CM334 287 ACAACCAGCATACTATACCAACCATATAATTTTAGGAAAAGATTCAACAGAATATAATA

CaBs4C.1 ECW+30R 371 TGGAGTTTGATCTAAGATACTTTCATCTTAATTTTGGCTAACATGCTCAAATCAATATTAT
CaBs4C.1 CM334 347 TGGAGTTTGATCTAAGATACTTTCATCTTAATTTTGGCTAACATGCTCAAATCAATATTAT

CaBs4C.1 ECW+30R 431 CCATTTCTAATAACTGGGATCCTTTCCATATAATTGGTAACCCCATTCAGTTTCATGA
CaBs4C.1 CM334 407 CCATTTCTAATAACTGGGATCCTTTCCATATAATTGGTAACCCCATTCAGTTTCATGA

CaBs4C.1 ECW+30R 491 CATTATGAATAGGCTCTTCTTTCTTTTCATATTTTCCTTCATTTTCTCCATCACTCGTA
CaBs4C.1 CM334 467 CATTATGAATAGGCTCTTCTTTCTTTTCATATTTTCCTTCATTTTCTCCATCACTCGTA

CaBs4C.1 ECW+30R 551 TGACACTTCATCATCCAATTATACGAATACGTGTAAGTATTACTTTCTCCAAGTCGTTTA
CaBs4C.1 CM334 527 TGACACTTCATCATCCAATTATACGAATACGTGTAAGTATTACTTTCTCCAAGTCGTTTA

```

<i>CaBs4C.1</i> ECW+30R	611	ATATCTTGTGCCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCTTCA
<i>CaBs4C.1</i> CM334	587	ATATCTTGTGCCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCTTCA
<i>CaBs4C.1</i> ECW+30R	671	TTATCATTACACTTTCCTTATGTTCTTCTTGGATTGGCAATACATGGGCTAGTTTTTCGAC
<i>CaBs4C.1</i> CM334	647	TTATCATTACACTTTCCTTATGTTCTTCTTGGATTGGCAATACATGGGCTAGTTTTTCGAC
<i>CaBs4C.1</i> ECW+30R	731	AACGGATTCTGCATATTTTCTCAACAATAATATTCCAGCAGTTAGCATTTTCATCAATG
<i>CaBs4C.1</i> CM334	707	AACGGATTCTGCATATTTTCTCAACAATAATATTCCAGCAGTTAGCATTTTCATCAATG
<i>CaBs4C.1</i> ECW+30R	791	TGGAGTTTGATAGAAATAATGTCATCAGCCAGCAGCATGAAACTCAAACAATCACCTGC
<i>CaBs4C.1</i> CM334	767	TGGAGTTTGATAGAAATAATGTCATCAGCCAGCAGCATGAAACTCAAACAATCACCTGC
<i>CaBs4C.1</i> ECW+30R	851	TCTCTCTCGTCTAGAAATCTAGATGAATTCATCGATCAAGTATTTTATGTTTCATCCAACGT
<i>CaBs4C.1</i> CM334	827	TCTCTCTCGTCTAGAAATCTAGATGAATTCATCGATCAAGTATTTTATGTTTCATCCAACGT
<i>CaBs4C.1</i> ECW+30R	911	CTTACATTTTCATATATGTTGTTTCCCTATGAAAATCTTGAAAATAAGGTAC-----
<i>CaBs4C.1</i> CM334	887	CTTACATTTTCATATATGTTGTTTCCCTATGAAAATCTTGAAAATAAGGTACGTTTAC

## A 24

Alignment der *CaBs4C.2*-Sequenzen aus *C. annuum* cv. CM334 (*CaBs4C.2* CM334) und aus *C. annuum* ECW+ECW-30R (*CaBs4C.2* ECW+30R). Grün markiert ist das Startkodon, rot das Stoppkodon und das *EBE*<sub>*CaBs4C.2*</sub> ist gelb hinterlegt.

Beide Sequenzen sind identisch. Grau markiert ist die Duplikation von 78 nt, die sowohl in CM334 als auch in ECW+ECW-30R identifiziert werden konnten.

### >*CaBs4C.2* (ECW+30R)

ATTATTAAGCGATAGAAACAATTAAAAATGCCAAAAATAGCCCTCCAATGCTAGTAGACCAGCCATGACAACTATTGTAAAGCCGTT  
CAATCTCGAGAAAAATTGGGCCAACCTATTATACACGCTCAACTTCGAGGAAAGAGCAGAATGTTAGAAATATAATTTTACGCG  
TAGATATAAATAATGAAAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCCACTCAAAATTTTCAAAAAAGCTCAAGA  
AGAAATGAGTTTGATCTTAGATACTTGATCTTAATTATAGTTAAACATTCTCAAATCCATATCACCTTCTGAAAATGGGATCCTTT  
CCATATATTTGGCAATCCCATTCAGTTTCATCACATTTATGAATAGGCTCGCCTTTCTTTTCTTATTTTCTTTCATTTTCTCCA  
TCACTCGTACAACTTCATCATCCAATTATACGAATACGTATAAGTATTACTTTCTCCAAGTCCCTTCATTTTCTCCATCACTCGT  
ACAACACTTCATCATCCAATTATACGAATACGTATAAGTATTACTTTCTCCAAGTCGTTTAAATATCTTGTGCCTAGCTTCTCTTCT  
ACTCCACAAAATGTTGTTCTGTAATTTTCTTCAATACCACTTTCCTCATGTTCTTCTTGGATTGGCAATGCATGGGCTA  
GTTTTCGACAACCGACTCTGCATATTTTCTTCAATATTTCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAAATACT  
GTGAGCCAGCAGCATGAAACTCAAACAACCTCGCCTGCTCTCTCTCGTCTAGACGAATTCATCGATCAAGGATATATTTCTTGTTC  
TCAACGTC

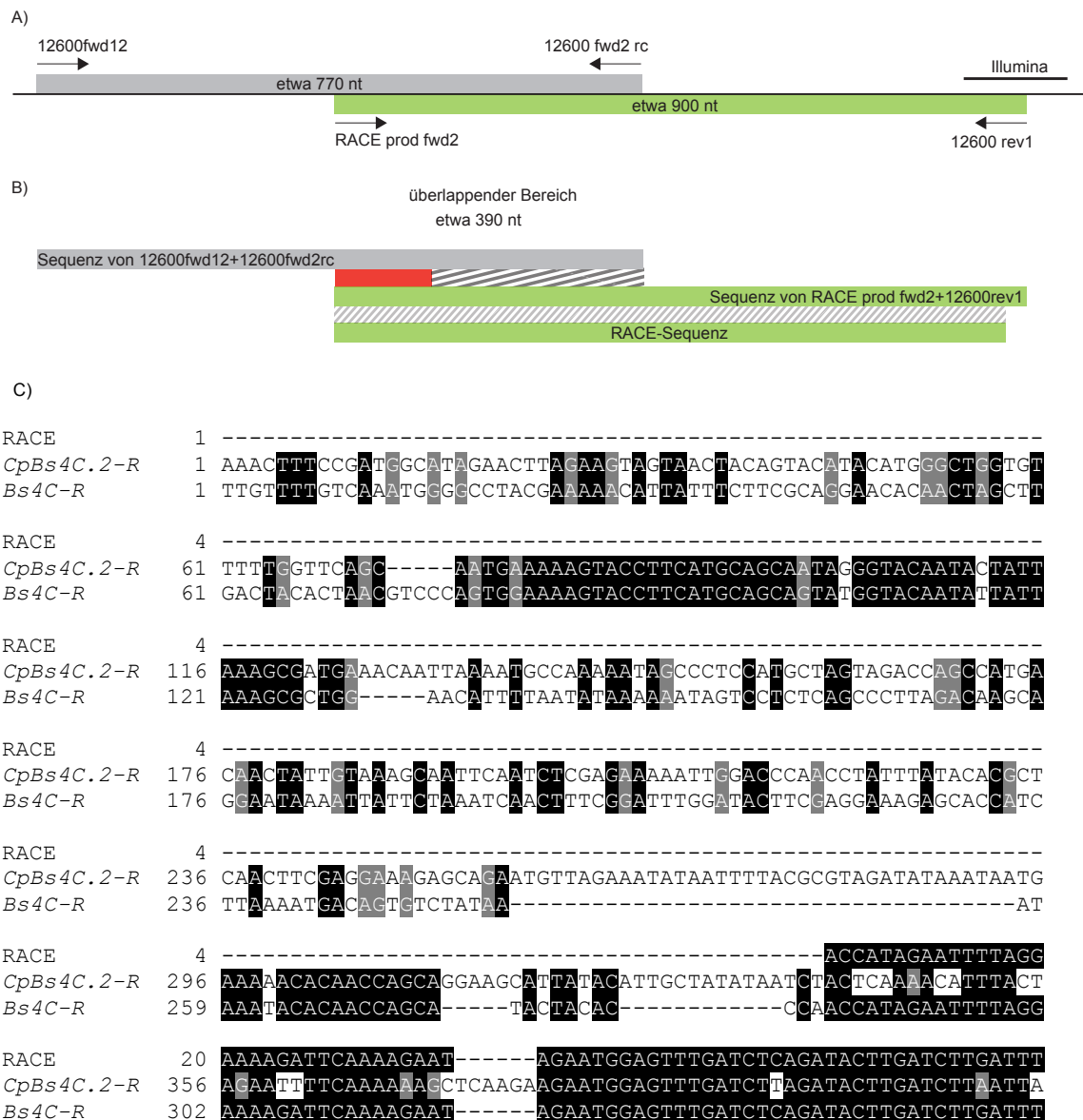
<i>CaBs4C.2</i> ECW+30R	1	ATTATTAAGCGATAGAAACAATTAAAAATGCCAAAAATAGCCCTCCAATGCTAGTAGACCA
<i>CaBs4C.2</i> CM334	1	-----GAAACAATTAAAAATGCCAAAAATAGCCCTCCAATGCTAGTAGACCA
<i>CaBs4C.2</i> ECW+30R	61	GCCATGACAACCTATTGTAAAGCCGTTCAATCTCGAGAAAAATTGGGCCAACCTATTTAT
<i>CaBs4C.2</i> CM334	46	GCCATGACAACCTATTGTAAAGCCGTTCAATCTCGAGAAAAATTGGGCCAACCTATTTAT
<i>CaBs4C.2</i> ECW+30R	121	ACACGCTCAACTTCGAGGAAAGAGCAGAATGTTAGAAATATAATTTTACGCGTAGATATA
<i>CaBs4C.2</i> CM334	106	ACACGCTCAACTTCGAGGAAAGAGCAGAATGTTAGAAATATAATTTTACGCGTAGATATA
<i>CaBs4C.2</i> ECW+30R	181	AATAATGAAAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCCACTCAAAAT
<i>CaBs4C.2</i> CM334	166	AATAATGAAAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCCACTCAAAAT
<i>CaBs4C.2</i> ECW+30R	241	TTTCAAAAAAGCTCAAGAAGAATGAGTTTGATCTTAGATACTTGATCTTAATTATAGTT
<i>CaBs4C.2</i> CM334	226	TTTCAAAAAAGCTCAAGAAGAATGAGTTTGATCTTAGATACTTGATCTTAATTATAGTT
<i>CaBs4C.2</i> ECW+30R	301	AACATTCTCAAATCCATATCACCTTCTGAAAATTTGGGATCCTTTCCATATATTTGGCAAT
<i>CaBs4C.2</i> CM334	286	AACATTCTCAAATCCATATCACCTTCTGAAAATTTGGGATCCTTTCCATATATTTGGCAAT
<i>CaBs4C.2</i> ECW+30R	361	CCCATTCCAAGTTTCATCACATTTATGAATAGGCTCGCCTTTCTTTTCTTATTTTCCTTC
<i>CaBs4C.2</i> CM334	346	CCCATTCCAAGTTTCATCACATTTATGAATAGGCTCGCCTTTCTTTTCTTATTTTCCTTC
<i>CaBs4C.2</i> ECW+30R	421	ATTTTCTCCATCACTCGTACAACACTTCATCATCCAATTATACGAATACGTATAAGTATT
<i>CaBs4C.2</i> CM334	406	ATTTTCTCCATCACTCGTACAACACTTCATCATCCAATTATACGAATACGTATAAGTATT
<i>CaBs4C.2</i> ECW+30R	481	ACTTTCTCCAAGTCCTTCATTTTCTCCATCACTCGTACAACACTTCATCATCCAATTATA
<i>CaBs4C.2</i> CM334	466	ACTTTCTCCAAGTCCTTCATTTTCTCCATCACTCGTACAACACTTCATCATCCAATTATA
<i>CaBs4C.2</i> ECW+30R	541	CGAATACGTATAAGTATTACTTTTCTCCAAGTCGTTTAAATATCTTGTGCCTAGCTTCTCTT
<i>CaBs4C.2</i> CM334	526	CGAATACGTATAAGTATTACTTTTCTCCAAGTCGTTTAAATATCTTGTGCCTAGCTTCTCTT
<i>CaBs4C.2</i> ECW+30R	601	CTACTCCCACAAATGTTGTTCTGGTACTTTTTCTTCATCACCATTTACACTTTCTCATGT
<i>CaBs4C.2</i> CM334	586	CTACTCCCACAAATGTTGTTCTGGTACTTTTTCTTCATCACCATTTACACTTTCTCATGT

<i>CaBs4C.2</i> ECW+30R	661	TCTTCTTGGATTGGCAATGCATGGGCTAGTTTTTCGACAACCGACTCTGCATATTTTCTCT
<i>CaBs4C.2</i> CM334	646	TCTTCTTGGATTGGCAATGCATGGGCTAGTTTTTCGACAACCGACTCTGCATATTTTCTCT
<i>CaBs4C.2</i> ECW+30R	721	ACAATATTTCCAGCAGTTAGCATTTTTCATCAATGTGGAGTTTGACAGAAATACTGTGAGC
<i>CaBs4C.2</i> CM334	706	ACAATATTTCCAGCAGTTAGCATTTTTCATCAATGTGGAGTTTGACAGAAATACTGTGAGC
<i>CaBs4C.2</i> ECW+30R	781	CAGCAGCATGAAACTCAAACAACCTCGCCTGCTCTCTCTCGTC <b>TAG</b> ACGAATTCATCGATC
<i>CaBs4C.2</i> CM334	766	CAGCAGCATGAAACTCAAACAACCTCGCCTGCTCTCTCTCGTC <b>TAG</b> ACGAATTCATCGATC
<i>CaBs4C.2</i> ECW+30R	841	AAGGATATATTTCTTGTTCATCAACGTC-----
<i>CaBs4C.2</i> CM334	826	AAGGATATATTTCTTGTTCATCAACGTCCTTGATTTTCATATATGTGTGTTTGTCTGTA

## A 25

Zur Bestimmung von zusätzlichen 5'Sequenzen wurde auf genomischem *C. pubescens*-Material mit folgenden Oligonukleotiden: RACE prod fwd2 und 12600rev1 sowie 12600 fwd12 und 12600fwd2rc amplifiziert und sequenziert (A). Das Alignment der erhaltenen Sequenzen mit der durch RACE-Analysen für *Bs4C-R* Sequenz zeigte Unterschiede im Überlappungsbereich und ließ vermuten, dass zwei *Bs4C*-Paraloge in *C. pubescens* (*CpBs4C.2-R*) vorhanden sind (B).

Das Alignment (C) wurde mit dem Programm T-coffee (<http://tcoffee.crg.cat/apps/tcoffee/index.html>) und Boxshade erstellt.



RACE	74	TGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTTCCATATAT
<i>CpBs4C.2-R</i>	416	TTTAAATCTCTCAAATCCATATCACC---TTCTGAAAAATTTGATCCTTTCCATATAT
<i>Bs4C-R</i>	356	TGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTTCCATATAT
RACE	134	TTCATGA--CCATCCC--AGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATAT
<i>CpBs4C.2-R</i>	473	TTGCAATCCATTCCAAGTTTCATCA--CATTATGAATAAGCTCGCCTTTGTTTTCATAT
<i>Bs4C-R</i>	416	TTCATGA--CCATCCC--AGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATAT
RACE	191	TTTCCTTTATTTTCTCCATCACTCGTATAAACACTTCATCATCCAAATATACGAATACGTG
<i>CpBs4C.2-R</i>	533	TTTCCTTTCATTTTCTCCATCACTCGTATAAACACTTCATCATCCAAATATACGAATACGTG
<i>Bs4C-R</i>	473	TTTCCTTTATTTTCTCCATCACTCGTATAAACACTTCATCATCCAAATATACGAATACGTG
RACE	251	TACGTACTACTACTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTGTGTCTAGCTTCTC
<i>CpBs4C.2-R</i>	593	TAGTATAC-----TTTCTCCAAGTCGTTAATAAGCTTGTGTCTAGCTTCTC
<i>Bs4C-R</i>	533	TACGTACTACTACTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTGTGTCTAGCTTCTC
RACE	311	TTCTACTCCACAAATGTTGTTCTGGTACTTTTTCGTCATTATCATTGCACCTTTCCCTCAT
<i>CpBs4C.2-R</i>	641	TTCTACTCCACAAATGTTGTTCTGGTACTTTTTCTCATTATCATTACACTTTCCCTCAT
<i>Bs4C-R</i>	593	TTCTACTCCACAAATGTTGTTCTGGTACTTTTTCGTCATTATCATTGCACCTTTCCCTCAT
RACE	371	GTTCTTCTTGGATTGGCAATACATGGGCTAGTTTTCGACAACGGATTCTGCATATTTTCT
<i>CpBs4C.2-R</i>	701	GTTCTTCTTGGATTAGC--CATACATGGGCTAGTTTCGACAACCGACTCTACATATTTTCT
<i>Bs4C-R</i>	653	GTTCTTCTTGGATTGGCAATACATGGGCTAGTTTTCGACAACGGATTCTGCATATTTTCT
RACE	431	CAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATGTGGAGTTTGACAGAAATAATG
<i>CpBs4C.2-R</i>	761	CTACAAT---ATTCCAGCAGTTAGCATTTTTCATCAATGTGGAGTTTGACAGAAATAATG
<i>Bs4C-R</i>	713	CAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATGTGGAGTTTGACAGAAATAATG
RACE	491	TGAGCCAGCAACATGAAACTCAAACAAATCGCCTGCTCGTCTAGTTGAATTCATCGATCA
<i>CpBs4C.2-R</i>	818	TGAGCCAGCAGCATGAAACTCAAACAAATCGCCTGCTCGTCAAG-----
<i>Bs4C-R</i>	773	TGAGCCAGCAACATGAAACTCAAACAAATCGCCTGCTCGTCTAGTTGAATTCATCGATCA
RACE	551	AGGATATATTTTCATGTTTCATCCAACGCTCTTGTATTTTCATATATGTGTGTTTTTCTGTA
<i>CpBs4C.2-R</i>	860	-----
<i>Bs4C-R</i>	833	AGGATATATTTTCATGTTTCATCCAACGCTCTTGTATTTTCATATATGTGTGTTTTTCTGTA

## A 26

Sequenzen von *CpBs4C.2-R* (PI 235047) und *CpBs4C.2-S* (PI 585270)Grün markiert ist das Startkodon, rot das Startkodon und die potentiellen *EBE* sind gelb hinterlegt.>*CpBs4C.2-R*

AAACTTTCCGATGGCATAGAACTTAGAAGTAGTAACACAGTACATACATGGGCTGGTGTGTTTTGGTTTCAGCAATGAAAAAGTACC  
 TTCATGCAGCAATAGGGTACAATACTATTAAAGCGATGAAACAATTAAAAAGCCAAATAAGCCCTCCATGCTAGTAGACAGCCA  
 TGACAACTATTGTAAGCAATTCAATCTCGAGAAAAATGGACCAACCTATTTATACAGCTCAACTTCGAGGAAAGAGCAGAAT  
 GTTAGAAATATAATTTACGCGTAGATATAAATAATGAAAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCTACTCA  
 AAACATTTACTAGAATTTTCAAAAAAGCTCAAGAAGAATGGAGTTTGATCTTAGATACTTGATCTTAATTATATTTAACATTCTCA  
 AATCATATCACCTTCTGAAAAATTTGATCCTTTCCATATATTTGGCAATCCCATTTCCAAGTTTCATCACATTTATGAATAGGCTC  
 GCCTTTGTTTTCTTATTTTCTTCCATTTTCTCCATCACTCGTATAAACACTTCATCATCCAATTATACGAATACGTATAAGTATTAC  
 TTTCTCCAAGTCGTTAATAGCTTGTGTCTAGCTTCTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTTCAATTATCATTA  
 CACTTTCTCATGTTCTTCTTGGATTAGCCATACATGGGCTAGTGTTCGACAACCGACTCTACATATTTTCTCTACAATATTGCCA  
 GCAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAAATAATGTGAGCCAGCAGCATGAAACTCCAACAAATCGCCTGCTCTCTCT  
 CGTCTAGATGAATCGATCGATCAAGGATTTTCATATTTATCCAACGCTTTACATTTTCATATATGTGTGTTTTGTTCTGTATGAAAT  
 CTTGAAATAAGGTATGTTTACCTACTATAACAAGTAAAGATGACACGATAGCGT

>*CpBs4C.2-S*

GGACAACTTTCCGATGGCATAGAACTTAGAAGTAGTAACACAGTACATACATGGGCTGGTGTGTTTTGGTTTCAGCAATGAAAAAG  
 TACCTTCATGCAGCAATAGGGTACAATACTATTAAAGCGTGAAACAATTAAAAAGCCAAATAAGCCCTCCATGCTAGTAGACCA  
 GCCATGACAACTATTGTAAGCAATTCAATCTCGAGAAAAATGGACCAACCTATTTATACGCTCAACTTCGAGGAAAGAGCAGA  
 ATGTTAGAAATATAATTTACGCGTAGATATAAATAATGAAAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCTACT  
 CAAAACATTTACTAGAATTTCTCAAAAAAGCTCAAGAAGAATGGAGTTTGATCTTAGATACTTGATCTTAATTATAGTTAACATTCT  
 CAAATCCATATCACCTTCTGAAAAATTTGATCCTTTCCATATATTTGGCAATCCCATTTCCAAGTTTCATCACATTTATGAATAGGC  
 TCGCCTTTGTTTTCTTATTTTCTTCCATTTTCTCCATCACTCGTATAAACACTTCATCATCCAATTATACGAATACGTATAAGTATT  
 ACTTTCTCCAAGTCGTTTAAATATCTGTGTCTAGCTTCTCTCTACTCCCAAAATGTTGTTCTGGTACTTTTCTTCAATTATCAT  
 TACACTTTCTCATGTTCTTCTTGGATTAGCCATACATGGGCTACTGTTTCGACAACCGACTCTGCATATTTTCTCTACAATATTGC  
 CAGCAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCCAACAAATCGCCTGCTCTCT  
 CTCGTCAGATGAATCGATCGATCAAGGATTTTCATATTTATCCAACGCTTTACATTTTCATATATGTGTGTTTTGTTCTGTATGAA  
 ATCTTGAAATAAGGTATGTTTACCTACTATAACAAGTAAAGATGACACGATAGC

Alignment der *CpBs4C.2-R* (PI 235047) und *CpBs4C.2-S* (PI 585270) Sequenzen (oben) und der Proteinsequenzen (unten). Grün markiert ist das Startkodon und die potentiellen *EBE* sind gelb hinterlegt. Identische Nukleotide zwischen beiden Akzessionen sind als weiße Buchstaben auf schwarzem Hintergrund dargestellt. Das Alignment wurde mit ClustalW und Boxshade erstellt.

```

CpBs4C.2-R   1  ---AAACTTTCCGATGGCATAGAACTTAGAAGTAGTAACTACAGTACATACATGGGCTG
CpBs4C.2-S   1  GGACAAACTTTCCGATGGCATAGAACTTAGAAGTAGTAACTACAGTACATACATGGGCTG

CpBs4C.2-R  57  GTGTTTTTGGTTCAGCAATGAAAAAGTACCTTCATGCAGCAATAGGGTACAATACTATTA
CpBs4C.2-S  61  GTGTTTTTGGTTCAGCAATGAAAAAGTACCTTCATGCAGCAATAGGGTACAATACTATTA

CpBs4C.2-R 117  AAGCGATGAAACAATTAAAAATGCCAAAAATAGCCCTCCATGCTAGTAGACCAGCCATGAC
CpBs4C.2-S 121  AAGCGCTGAAACAATTAAAAATGCCAAAAATAGCCCTCCATGCTAGTAGACCAGCCATGAC

CpBs4C.2-R 177  AACTATTGTAAAGCAATTCAATCTCGAGAAAAATTGGACCCAACCTATTTATACACGCTC
CpBs4C.2-S 181  AACTATTGTAAAGCAATTCAATCTCGAGAAAAATTGGACCCAACCTATTTATAC--GCTC

CpBs4C.2-R 237  AACTTCGAGGAAAGAGCAGAATGTTAGAAATATAATTTTACGCGTAGATATAAATAATGA
CpBs4C.2-S 239  AACTTCGAGGAAAGAGCAGAATGTTAGAAATATAATTTTACGCGTAGATATAAATAATGA

CpBs4C.2-R 297  AAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCTACTCAAAACATTTACTA
CpBs4C.2-S 299  AAAACACAACCAGCAGGAAGCATTATACATTGCTATATAATCTACTCAAAACATTTACTA

CpBs4C.2-R 357  GAATTTTCAAAAAAGCTCAAGAAGAATGAGAGTTTGATCTTAGATACTTGATCTTAATTAT
CpBs4C.2-S 359  GAATTTTCAAAAAAGCTCAAGAAGAATGAGAGTTTGATCTTAGATACTTGATCTTAATTAT

CpBs4C.2-R 417  ATTAAACATTCTCAATCCATATCACCTTCTGAAAATTTTGATCCTTTCCATATATTTGG
CpBs4C.2-S 419  AGTTAAACATTCTCAATCCATATCACCTTCTGAAAATTTTGATCCTTTCCATATATTTGG

CpBs4C.2-R 477  CAATCCCATTCCAAGTTTCATCACATTTATGAATAGGCTCGCCTTTGTTTTCTTATTTTC
CpBs4C.2-S 479  CAATCCCATTCCAAGTTTCATCACATTTATGAATAGGCTCGCCTTTGTTTTCTTATTTTC

CpBs4C.2-R 537  CTTCAATTTTCTCCATCACTCGTATAACACTTCATCATCCAATTATACGAATACGTATAAG
CpBs4C.2-S 539  CTTCAATTTTCTCCATCACTCGTATAACACTTCATCATCCAATTATACGAATACGTATAAG

CpBs4C.2-R 597  TATTACTTTTCTCCAAGTCGTTTAATAGCTTGTGTCTAGCTTCTCTTCTACTCCACAAAT
CpBs4C.2-S 599  TATTACTTTTCTCCAAGTCGTTTAATAGCTTGTGTCTAGCTTCTCTTCTACTCCACAAAT

CpBs4C.2-R 657  GTTGTTCTGGTACTTTTCTTCATTATCATTACACTTTCCTCATGTTCTTCTTGGATTAG
CpBs4C.2-S 659  GTTGTTCTGGTACTTTTCTTCATTATCATTACACTTTCCTCATGTTCTTCTTGGATTAG

CpBs4C.2-R 717  CCATACATGGGCTAGTGTTCGACAACCGACTCTACATATTTTCTCTACAATATTGCCAGC
CpBs4C.2-S 719  CCATACATGGGCTAGTGTTCGACAACCGACTCTGCATATTTTCTCTACAATATTGCCAGC

CpBs4C.2-R 777  AGTTAGCATTTTTCATCAATGTGGAGTTTGACAGAAATAATGTGAGCCAGCAGCATGAAAC
CpBs4C.2-S 779  AGTTAGCATTTTTCATCAATGTGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAAC

CpBs4C.2-R 837  TCCAACAAATCGCCTGCTCTCTCTCGTCTAGATGAATCGATCGATCAAGGATTTTCATATT
CpBs4C.2-S 839  TCCAACAAATCGCCTGCTCTCTCTCGTCTAGATGAATCGATCGATCAAGGATTTTCATATT

CpBs4C.2-R 897  TATCCAACGTCTTACATTTTCATATATGTGTGTTTGTCTGTATGAAAATCTTGAAAATA
CpBs4C.2-S 899  TATCCAACGTCTTACATTTTCATATATGTGTGTTTGTCTGTATGAAAATCTTGAAAATA

CpBs4C.2-R 957  AGGTATGTTTACCTACTATAACAAGTAAAAGATGACACGATAGCGT
CpBs4C.2-S 959  AGGTATGTTTACCTACTATAACAAGTAAAAGATGACACGATAGC--

```

#### >CpBs4C.2-R

```

MEFDRLYLILIIIFNILKSISPSENFDPFHIFGNPIPSFITFMNRLAFVFLFSFIFSIITRITLHHPIIRIRISITFS
KSFNSLCLASLLLPQMLFWYFFFIITLSSCSSLWISHTWASVRQPTLHIFSTILPAVSIFINVEFDRNNVSQQHET
PTNRLLSLV

```

#### >CpBs4C.2-S

```

MEFDRLYLILIIIVNILKSISPSENFDPFHIFGNPIPSFITFMNRLAFVFLFSFIFSIITRITLHHPIIRIRISITFS
KSFNILCLASLLLPQMLFWYFFFIITLSSCSSLWISHTWATVRQPTLHIFSTILPAVSIFINVEFDRNNVSQQHET
PTNRLLSLV

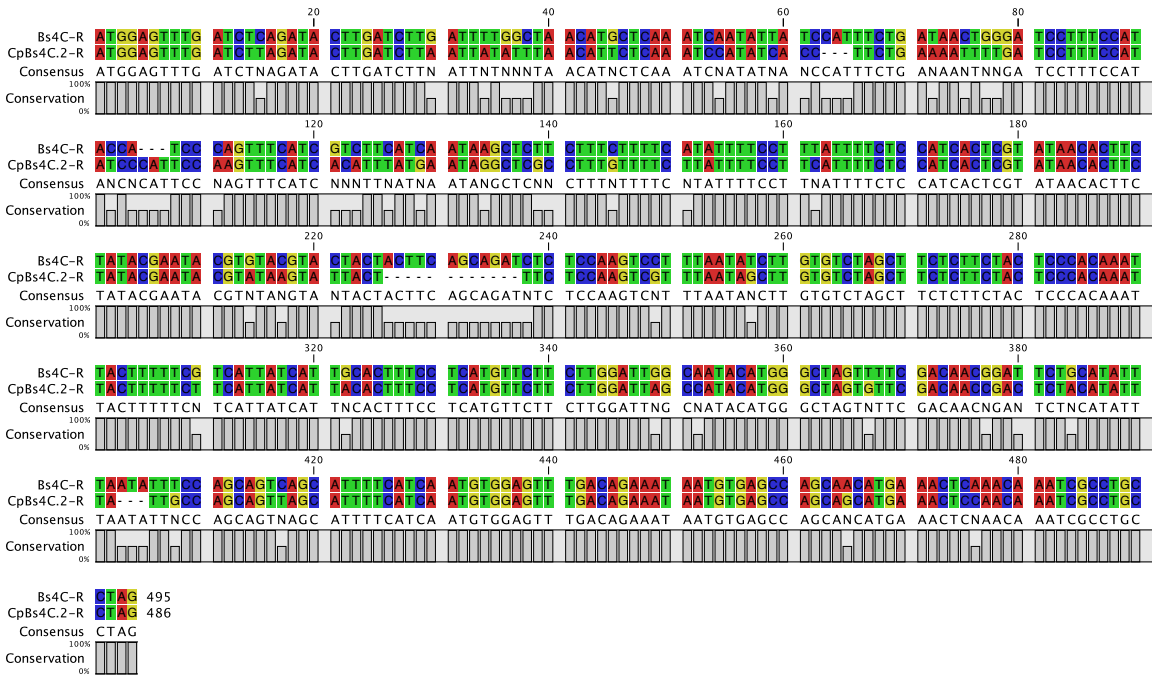
```

Vergleich der CpBs4C.2-R (PI 235047) und CpBs4C.2-S (PI 585270) Sequenzen

CpBs4C.2-R	1	MEFDRLYLILII	FNILKSISPS	ENFDPFHIFGN	PIPSFITFMNRLAFVFLFSFIFSITRI
CpBs4C.2-S	1	MEFDRLYLILII	FNILKSISPS	ENFDPFHIFGN	PIPSFITFMNRLAFVFLFSFIFSITRI
CpBs4C.2-R	61	TLHHPIIRIRISITFSKSFNS	LCLASLLLPQMLFWYFFFI	IIITLSSCSSWISHTWASVRQ	
CpBs4C.2-S	61	TLHHPIIRIRISITFSKSFNI	LCLASLLLPQMLFWYFFFI	IIITLSSCSSWISHTWATVRQ	
CpBs4C.2-R	121	PTLHIFSTILPAVSIFINVEFDRN	NVVSQQHETPTNRLLSLV		
CpBs4C.2-S	121	PTLHIFSTILPAVSIFINVEFDRN	NVVSQQHETPTNRLLSLV		

A 27

Alignment und Vergleich der CpBs4C.2-R (PI 235047) und Bs4C-R Sequenzen



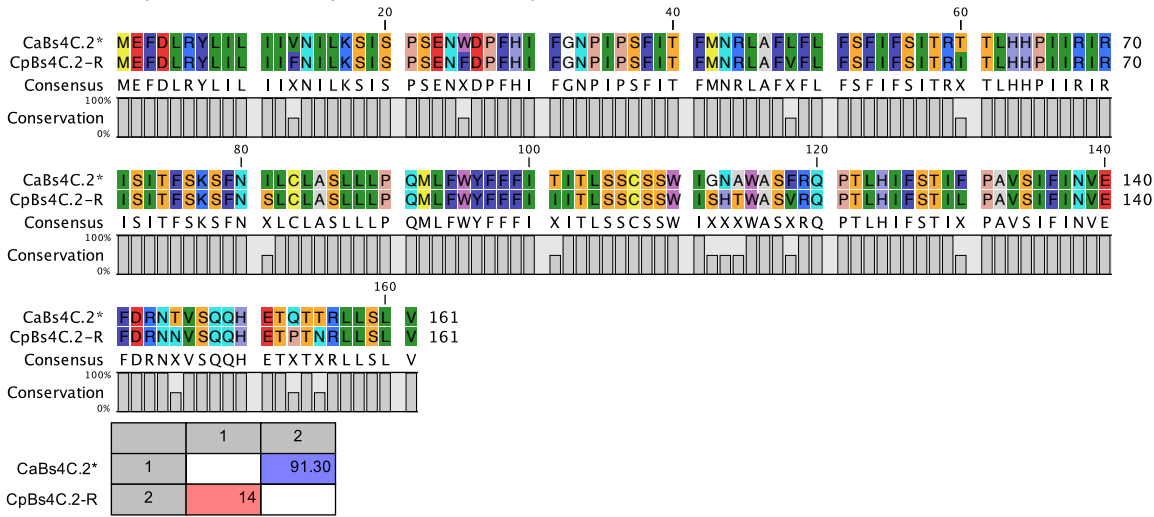
Alignment und Vergleich der CpBs4C.2-R (PI 235047) und Bs4C-R Sequenzen





A 30

Alignment und Vergleich von CpBs4C.2-R (PI 235047) mit dem Deletionsderivat von CaBs4C.2. (CaBs4C.2\*, Duplikation deletiert)





**A 31**Sequenzen ATG-STOPP der verschiedenen *Capsicum*-Spezies

&gt;Bs4C-R

ATGGAGTTTGTATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCTTTAATATCTTG  
TGCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCGTATTATCATTGCACCTTCCTCATGTTCTCTTGGAT  
TGGCAATACATGGGCTAGTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATTTCCAGCAGTCAGCATTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAAATCGCCTGCTCGTCTAG

&gt;Bs4C-S

ATGGAGTTTGTATCTCAGATACTTCATCTTAATTTTGGCTAACATGATCAAATCAATATTATCATCCATTTCTGATAACTGGGATCC  
TTTCCATATATTTTCATGACCATCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA  
TCACTCGTATAACACTTTCATCATCCAAATATACGAATACGTATACGTACTACTTTCAGCAGATCTCTCCAAGTCTTTAATATC  
TTGTGCCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCATCACCATTACACTTTCTCATGTTCTCTTGG  
GATTGGCAATGCATGGGCTAGTTTTCGAGGAAGGACTCTGCATATTTTCTCTACAATATTTCCAGCAGTTAGCATTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAAATCGCCTGCTCGTCTAG

&gt;CpBs4C.2-R

ATGGAGTTTGTATCTTAGATACTTGATCTTAATTATATTTAACATTCTCAAATCCATATCACCTTCTGAAAATTTTGATCCTTTCCA  
TATATTTGGCAATCCCATTTCCAAGTTTCATCACATTTATGAATAGGCTCGCCTTTGTTTTCTTATTTTCTTCATTTTCTCCATCA  
CTCGTATAACACTTTCATCATCCAAATATACGAATACGTATAAGTATTACTTTCTCCAAGTCGTTTAATAGCTTGTGTCTAGCTTCT  
CTTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCTTCATTATCATTACACTTTCTCATGTTCTCTTGGATTAGCCATACATG  
GGCTAGTGTTTCGACAACCGACTCTACATATTTTCTCTACAATATTGCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAA  
ATAATGTGAGCCAGCAGCATGAAACTCCAACAAATCGCCTGCTCTCTCTCGTCTAG

&gt;CpBs4C.2-S

ATGGAGTTTGTATCTTAGATACTTGATCTTAATTATAGTTAACATTCTCAAATCCATATCACCTTCTGAAAATTTTGATCCTTTCCA  
TATATTTGGCAATCCCATTTCCAAGTTTCATCACATTTATGAATAGGCTCGCCTTTGTTTTCTTATTTTCTTCATTTTCTCCATCA  
CTCGTATAACACTTTCATCATCCAAATATACGAATACGTATAAGTATTACTTTCTCCAAGTCGTTTAATATCTTGTGTCTAGCTTCT  
CTTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCTTCATTATCATTACACTTTCTCATGTTCTCTTGGATTAGCCATACATG  
GGCTACTGTTTCGACAACCGACTCTGCATATTTTCTCTACAATATTGCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAA  
ATAATGTGAGCCAGCAACATGAAACTCCAACAAATCGCCTGCTCTCTCTCGTCTAG

&gt;CaBs4C.1

ATGGAGTTTGTATCTAAGATACTTCATCTTAATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTAATAACTGGGATCCTTT  
CCATATAATTTGGTAACCCCATTTCCAAGTTTCATGACATTTATGAATAGGCTCTCTTTCTTTTCATATTTTCTTCATTTTCTCCA  
TCACTCGTATGACACTTTCATCATCCAATTATACGAATACGTGTAAGTATTACTTTCTCCAAGTCGTTTAATATCTTGTGTCTAGCT  
TCTCTTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCTTCATTATCATTACACTTTCTCATGTTCTCTTGGATTGGCAATAC  
ATGGGCTAGTTTTCGACAACCGATTCTGCATATTTTCTCTACAATAATATTTCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTG  
ATAGAAATAATGTATCATCAGCCAGCAGCATGAAACTCAAACAAATCACCTGCTCTCTCTCGTCTAG

&gt;CaBs4C.2

ATGGAGTTTGTATCTTAGATACTTGATCTTAATTATAGTTAACATTCTCAAATCCATATCACCTTCTGAAAATTTGGGATCCTTTCCA  
TATATTTGGCAATCCCATTTCCAAGTTTCATCACATTTATGAATAGGCTCGCCTTTCTTTTCTTATTTTCTTCATTTTCTCCATCA  
CTCGTACAACACTTTCATCATCCAATTATACGAATACGTATAAGTATTACTTTCTCCAAGTCCTTCATTTTCTCCATCACTCGTACA  
ACACTTCATCATCCAATTATACGAATACGTATAAGTATTACTTTCTCCAAGTCGTTTAATATCTTGTGTCTAGCTTCTCTTCTACT  
CCCACAAATGTTGTTCTGGTACTTTTTCTTCATCACCATTACACTTTCTCATGTTCTCTTGGATTGGCAATGCATGGGCTAGTT  
TTCGACAACCGACTCTGCATATTTTCTCTACAATAATTTCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTGACAGAAATACTGTG  
AGCCAGCAGCATGAAACTCAAACAACTCGCCTGCTCTCTCTCGTCTAG

&gt;CaBs4C.3

ATGGAGTTTGTATCTGATTATAGTTAACATACATTCTCAAATCCATATCACCTTCTGATCAAAATGTGATCCTTTCCATATATTT  
GGCAATCCCATTTCCAAGTTTCATGACATTTATGAATAGGCTCTCTTTCTCTCTATTTTCTCTTCATTTTCTCCATCACTCGTAT  
AACACTTCATCATCCAATTATACGAATACATATAAGTACCCTACTTTCAGCAGATTCTCCAAGTCTTTAATATCTTGTGTCTAG  
CTTCTCTTCTACTCCCACAAATGTTATTTTGGTACATTTTCTTCATCACCATTACATTGACCTCATGTTCTCTTGGGTTGGCAAT  
ACATGGGTTAGTTTTCGACAACGGACTTTGCGTATTTTCTCTACAATAATTTCCAGCAGTTAGCATTTTCATCAATGTGGAGTTTGA  
TAGAAATAACGTAAGCCAGCAGCTGAAACTCAAACAACTACTCATTTACTCTCTCTCTCGTCTAG

&gt;CaBs4C.4

ATGGAGTTTGTATCTTACATACTTTATCTTGATTATAGTTAACATACATTCTCAAATCCATGTCATCTTCAGAAAATTTGGGATCCT  
TTCCATATATTTGGCAGTCCCATTTCAAGTATCATGACATTTATGAACAGGCTCGTCTTTCTTTTCATATTTTCTTCATTTTCTC  
CATCACTCGTATGACACTTTCATCATCCAATTATACAAATACGTATAAGTATTACTTTCTCCAAGTCTTTAATATCTTGTGTCTAG  
CTTCTCTTCTACTCCCACAAATGTTGTTCTGGTACATTTTCTTCATTATCATTACATTTTCCCCATGTTCTCTTGGATTGGCAAT  
ACATGGGCTAGTTTTCGACAATGGACTTTGCGTATTTTCTCTACAATAATTTCCAGCAGTTTGCATTTTTCAGCAATTTGCACGT  
TGATAGATATTCATCATCTGCAGAAACATTTTCCATGTTAACTATTATGCTGAAGGCTACTGATGCTTTTTTAA

&gt;CaBs4C.5

ATGGAGTTTGTATCTAAGGTACACAATATTAATTATAGTTAACATTCTCAAATCAATATTTATTTCAAATGAATGGGATCCTTTTCA  
TATATTTGGCAACTCACTTTCATGGCATTATGAATAGGCTTTTCTCTTCTATTTGTTTCTTTTCATTTTCTCCATCACTAGTA  
TAACAATTCATTATCCAATGAGAGTACATATAAGTTCAACATATTTCTCAAAGTCTTTAATATATTTTAACTTCATTTCTA  
CTTCCACAAGTGTTATTTTGGTACATTTTCTTCATTACTATTATGACATTTCTTCTAATTGGATTCCAATTTTAAAGGAATGGTC  
TTTATATATTTTCAATAATTCAGGATTTAGTATTTTTCATCAATGCTGAGCTTGATAGAAATAATGTGAATAACCAGCAGCCCC  
AAGCTCAAACAATTCACCTACTCCTTCTCTGTCTAA

## &gt;CaBs4C.6

ATGGAGTTTTTAGAGGAAGTAGTGACGCACAGTCTCAGTTAAAAATTAGTGCTGTAAAAATATCACAAAGTGATTTTCATCTCTCACT  
TTTATATTGCAGATATGTAAGTTGCAATTTTCAAGGATTGATATGACGAAGGCATTTTCATTCTACTATCCAAATATTGTGTCATCCTT  
TGTTTACCCCATTTGAGCTTTAGTCTTATTTTATTTTCATACCCCTTTATTGGAATCAAGATAGAGTCGGTAAAGCTCAAAAGAAAA  
GTGTCGAGCAAAACAATAGAGTCAGAAAGTTTCAAAAGAAACGAGACAAAAAGTGTCTAAAAGAGAAAAAGAAAGCGTCGAGA  
AAAATAA

## &gt;CaBs4C.7

ATGGAGTTTAGAACTGGCCTAGGGTCGCTGCTTCTGATTGCCCTTGTGGAGTTCCTACTGGTTTCCTCTAGCCCTCTTCAGGCACCT  
TAAGTTGGCTTATGTAATACTGAGTTTTGACAGTTGATCCATTTGCTTACCCAGTTGGTGACTTCTCAAGCCGAGCATGCTAATTT  
TGTTGCTCCATCTTCTATGGTTAATAAGGTTGGCCAATTTATAAGGTTGAGTTCTCCAACCTTCTACTGGTTCTAAAGTCGAGGAAT  
ATCCTCAAAATTTTCATTGATGAGATAGAAAAAAATTTTCGTATAATGCATGCTATGATTGTAGGGTGTGGAGTTTGTGTCCTA  
TCAGTTGAATTATGTGACGTATCAGTGGTATTATAGGGACGATGATGACAGTCCACTCTATGGGATGACTTCTCTAGTGGTTTTCT  
TTGAATATTTTTTCTTCAGGAGTTGAGGGGTGAAGGCTAAGGAATTTGAGAATTTAAAGTAA

>CAP468P2 (*C. pubescens* CAP468 Pflanze 2)

ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTTCATCATCCAAATATACGAATACGTATACGTACTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGCTAGCTTCTCTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTCGTCATTATCATTGCACCTTTCCTCATGTTCTCTTGGAT  
TGGCAATACATGGGCTAGTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAATCGCCTGCTCGTCTAG

>*C. tovari* (*C. tovari* Eshba CGN22876)

ATGGAGTTTGATCTAAGATACTTGATCTTAAATTTTGGCTAACATGGTCAAATCAATATCATCATCCATTTCTGATAACTGGGATCC  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA  
TCACTCGTATAACACTTTCATCATCCAAATATACGGATACGTATACGTACTACTACTTTCAGCAGATCTCTCAAAGTCCTTTAATATC  
TTGTTCTAGCTTCTCTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTTCTTATCATTACACTTTTCCTCATGTTCTCTTGGAT  
GATTGGCAATGCATGGGCTAGTTTTCGACAACACTCTCTGCATATTTTCTCTACAATATTTCCAGCAGTTAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAATCGCCTGCTCTCTCTCGTCTAG

>*C. frutescens*

ATGGAGTTTGATCTAAGATACTTTCATCTTAAATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTTCATCATCCAAATATACGAATACGTATACGTACTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGCTAGCTTCTCTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTTTCATCACCATTACACTTTCCTCATGTTCTCTTGGAT  
TGGCAATGCATGGGCTAGTTTTCGACAACGGACTCTGCATATTTTCTCTACAATATTTCCAGCAGTTAGCATTTTTCATCAATGTGG  
AGTTTGACAGAAATACTGTGAGCCAGCAGCATGAAACTCAAACAATCGCCTGCTCTCTCTCGTCTAG

>*C. galapo* (*C. galapogense* CGN22208)

ATGGAGTTTGATCTCAGATACTTTCATCTTAAATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTAATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATGACATTTATGAATAGGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA  
TCACTCGTATGACACTTTCATCATCCAAATATACGAATACGTGTAAGTATTACTTTCTCCAAGTCGTTTAAATATCTTGTGCTAGCT  
TCTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTTTCATATCATTACACTTTCCTTATGTTCTTCTTGGATTGGCAATAC  
ATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTTAGCATTTTTCATCAATGTGGAGTTTG  
ATAGAAATAATGTCTATCAGCCAGCAGCATGAAACTCAAACAATCGCCTGCTCTCTCTCGTCAA

>*C. chaco* (*C. chacoense* PI560944)

ATGGAGTTTGATCTAAGATACTTTCATCTTAAATTTTGGCTAACATGCTCAAATCAATATTATCATCCATTTCTGATAACTGGGATCC  
TTTCCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA  
TCACTCGTATAACACTTTCATCATCCAAATATACGAATACGTATACGTACTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATC  
TTGTTCTAGCTTCTTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTTTCATCACCATTACACTTTCCTCATGTTCTCTTGGAT  
GATTGGCAATGCATGGGCTAGTTTTTCGACAACCGACTCTGCATATTTTCTCTACAATATTTCCAGCAGTTAGCATTTTTCATCAATG  
CGGAGTTTGACAGAAATACTGTGAGCCAGCAGCATGAAACTCAAACAATCGCCTGCTCTCTCTCGTCTAG

>*C. praeterX* (*C. praetermissum* CGN19198 Pflanze X)

ATGGAGTTTGATCTAAGATACTTTCATCTTAAATTTTGGCTAACATGATCAAATCAATATCATCATCCATTTCTGATAACTGGGATCC  
TTTCCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA  
TCACTCGTATAACACTTTCATCATCCAAATATACGAATACGTATACGTACTGCTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTTAGCTTCTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTTTCATCACCATTACACTTTCCTCATGTTCTCATCTTG  
GATTGGCAATGCATGGGCTAGTTTTTCGACAACCGACTCTGCATATTTTCTCTACAATATTTCCAGCAGTTAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAATCGCCTGCTCTCTCTCGTCTAG

>19243 (*C. pubescens* CGN19243)

ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTTCATCATCCAAATATACGAATACGTATACGTACTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTTAGCTTCTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTCGTCATTATCATTGCACCTTTCCTCATGTTCTCTTGGAT  
TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAATCGCCTGCTCTCTCGTCTAG

>1481-1 (*C. pubescens* CAP1481 Pflanze 1)

ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTTCATCATCCAAATATACGAATACGTATACGTACTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTTAGCTTCTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTCGTCATTATCATTGCACCTTTCCTCATGTTCTCTTGGAT  
TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAATCGCCTGCTCTCTCGTCTAG

TGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCGTCAATTATCATTGCACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAACTCAAACAAATCGCCTGCTCGTCTAG

>1482-4 (*C. pubescens* CAP1482 Pflanze 4)

ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCGTCAATTATCATTGCACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAACTCAAACAAATCGCCTGCTCGTCTAG

>1483-2 (*C. pubescens* CAP1483 Pflanze 2)

ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCGTCAATTATCATTGCACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAACTCAAACAAATCGCCTGCTCGTCTAG

>1492-1 (*C. pubescens* CAP1492 Pflanze 1)

ATGGAGTTTGATCTCAGATACTTCATCTTAATTTTGGCTAACATGATCAAATCAATATTATCATCCATTTCTGATAACTGGGATCC  
TTTCCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA  
CTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TTGTGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTCTTCATCACCATTACACTTTCCTCATGTTCTTCTTGGAT  
GATTGGCAATGCATGGGCTAGTTTTTCGAGGAAGGACTCTGCATATTTTCTTACAATATTTCCAGCAGTTAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAACTCAAACAAATCGCCTGCTCGTCTAG

>1486-1 (*C. pubescens* CAP1486 Pflanze 1)

ATGGAGTTTGATCTCAGATACTTCATCTTAATTTTGGCTAACATGATCAAATCAATATTATCATCCATTTCTGATAACTGGGATCC  
TTTCCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCA  
TCACTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TTGTGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTCTTCATCACCATTACACTTTCCTCATGTTCTTCTTGGAT  
GATTGGCAATGCATGGGCTAGTTTTTCGAGGAAGGACTCTGCATATTTTCTTACAATATTTCCAGCAGTTAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAACTCAAACAAATCGCCTGCTCGTCTAG

>870-1 (*C. pubescens* CAP879 Pflanze 1)

ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCGTCAATTATCATTGCACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAACTCAAACAAATCGCCTGCTCGTCTAG

>273-1 (*C. pubescens* CAP271 Pflanze 1)

ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTTCGTCAATTATCATTGCACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAACTCAAACAAATCGCCTGCTCGTCTAG

>*C.bac306-4* (*C. baccatum* 306 Pflanze 4)

ATGGAGTTTGATCTAAGATACTTCATCTTAATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTAATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTTATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTCTAGCTTCTCTCTACTCCCACAAATGTTGTTCTGGTACTTTTCTTCATCACCATTACACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATGCATGGGCTAGTTTTTCGACAACCGACTCTGCATATTTTCTTACAATATTTCCAGCAGTTAGCATTTTTCATCAATGTGG  
AGTTTGACAGAAATACTGTGAGCCAGCAGCATGAACTCAAACAACTCGCCTGCTCTCTCTCGTCTAG

>*C.bac874-6* (*C. baccatum* 874 Pflanze 6)

ATGGAGTTTGATCTAAGATACTTCATCTTAATTTTGGCTAACATGCTCAAATCAATATTATTCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATGACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTCTAGCTTCTTTTCTACTCCCACAAATGTTGTTCTGGTACTTTTCTTCATCACCATTACACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATGCATGGGCTAGTTTTTCGACAACCGACTCTGCATATTTTCTTACAATATTTCCAGCAGTTAGCATTTTTCATCAATGCGG  
AGTTTGACAGAAATACTGTGAGCCAGCAGCATGAACTCAAACAACTCGCCTGCTCTCTCTCGTCTAG

>1475-1 (*C. pubescens* CAP1475 Pflanze 1)

ATGGAGTTTGATCTAAGATACTTCATCTTAATTTTGGCTAACATGCTCAAATCAATATTATTCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACCATCCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCTTTATTTTCTCCATCA  
CTCGTATGACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATCTTG  
TGTCTAGCTTCTTTTCTACTCCCACAAATGTTGTTCTGGTACTTTTCTTCATCACCATTACACTTTCCTCATGTTCTTCTTGGAT  
TGGCAATGCATGGGCTAGTTTTTCGACAACCGACTCTGCATATTTTCTTACAATATTTCCAGCAGTTAGCATTTTTCATCAATGCGG  
AGTTTGACAGAAATACTGTGAGCCAGCAGCATGAACTCAAACAACTCGCCTGCTCTCTCTCGTCTAG

>1480-4 (*C. pubescens* CAP1480 Pflanze 4)  
ATGGAGTTTGATCTCAGATACTTGATCTTGATTTTGGCTAACATGCTCAAATCAATATTATCCATTTCTGATAACTGGGATCCTTT  
CCATATATTTTCATGACGATCCCGAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTCTTTTCATATTTTCCCTTTATTTTCTCCATCA  
CTCGTATAACACTTCATCATCCAAATATACGAATACGTGTACGTACTACTCTCAGCAGATCTCTCCAAGTCCCTTTAATATCTTG  
TGCTAGCTTCTCTTCTACTCCACAAATGTTGTTCTGGTACTTTTCGTGTCATTATCATTTGCACTTTCTCATGTTCTCTCTTGAT

TGGCAATACATGGGCTAGTTTTTCGACAACGGATTCTGCATATTTTCTCAACAATAATATTTCCAGCAGTCAGCATTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAAATCGCCTGCTCGTCTAG

>1480-2 (C. pubescens CAP1480 Pflanze 2)

ATGGAGTTTGATCTCAGATACTTCATCTTAATTTTGGCTAACATGATCAAATCAATATTATCATCCATTTCTGATAACTGGGATCC  
TTTCCATATATTTTCATGACCATCCAGTTTCATCGTCTTCATCAATAAGCTCTTCTTTTCATATTTTCCTTTATTTTCTCCA  
TCACTCGTATAACACTTCATCATCCAAATATACGAATACGTATACGTACTACTTTCAGCAGATCTCTCCAAGTCCTTTAATATC  
TTGTGTCTAGCTTCTCTTCTACTCCCAAAATGTTGTTCTGGTACTTTTTCTTCATCACCATTACACTTTCTCATGTTCTTCTTG  
GATTGGCAATGCATGGGCTAGTTTTTCGAGGAAGGACTCTGCATATTTTCTCTACAATATTTCCAGCAGTTAGCATTTTCATCAATG  
TGGAGTTTGACAGAAATAATGTGAGCCAGCAACATGAAACTCAAACAAATCGCCTGCTCGTCTAG

## Auflistung der korrespondierenden Proteinsequenzen zu den KDS-Sequenzen der verschiedenen *Capsicum*-Spezies.

>C.exi6-1

MEFDLRYFILILANMLKSILFISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRIRITTSADLSKSFNII  
CLASLLLPQMLFWYFFFITITLSSCSSWIGNAWASFRQPTLHIFSTIIFPAVSIFINVEFDRNTVSQQHETQTRLLSLV

>C.exi21502

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRIRITTSADLSKSFNII  
LASLLLPQMLFWYFFFITITLSSCSSWIGNAWASFRGRTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>1694-2

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRVRTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>867-11

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRVRTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>907-5

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRVRTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>C.bac306-4

MEFDLRYFILILANMLKSILSISNNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRISTTSADLSNSFNII  
CLASLLLPQMLFWYFFFITITLSSCSSWIGNAWASFRQPTLHIFSTIIFPAVSIFINVEFDRNTVSQQHETQTRLLSLV

>C.bac874-6

MEFDLRYFILILANMLKSILFISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRMTLHHPNIRIRIRITTSADLSKSFNII  
CLASFLLPQMLFWYFFFITITLSSCSSWIGNAWASFRQPTLHIFSTIIFPAVSIFINAEFDRNTVSQQHETQTRLLSLV

>1475-1

MEFDLRYFILILANMLKSILFISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRMTLHHPNIRIRIRITTSADLSKSFNII  
CLASFLLPQMLFWYFFFITITLSSCSSWIGNAWASFRQPTLHIFSTIIFPAVSIFINAEFDRNTVSQQHETQTRLLSLV

>217-5

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRVRTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>1483-2

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRVRTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>1492-1

MEFDLRYFILILANMIKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRIRITTSADLSKSFNI  
LCLASLLLPQMLFWYFFFITITLSSCSSWIGNAWASFRGRTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>1486-1

MEFDLRYFILILANMIKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRIRITTSADLSKSFNI  
LCLASLLLPQMLFWYFFFITITLSSCSSWIGNAWASFRGRTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>870-1

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRIRITTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>273-1

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRIRITTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>C.praeterX

MEFDLRYFILILANMIKSISISSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSI TRITLHHPNIRIRIRITASADLSKSFNII  
CLASLLLPQMLFWYFFFITITLSSCSHSWIGNAWASFRQPTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLV

>19243

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRTTTTADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>1481-1

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRVRTTTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>1482-4

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRVRTTTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>C.chaco

MEFDLRYFILILANMLKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRTTTTADLSNSFNI  
LCLASFLLPQMLFWYIFFITITLSSCSSWIGNAWASFRQPTLHIFSTIIFPAVSIFINAEFDRNTVSQQHETQTRLLSLV

>C.galapo

MEFDLRYFILILANMLKSILSISNNWDPFHIIIGNPIPSFMTFMNRLFFLFIFSFIFSITRMTLHHPNIRIRVSITFSKSFNII  
CLASLLLPQMLFWYFFFIITLSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVISQQHETQTNRLLV

>C.frut

MEFDLRYFILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRATTSADLSKSFNII  
CLASLLLPQMLFWYFFFIITLSSCSSWIGNAWASFRQRTLHIFSTIIFPAVSIFINVEFDRNTVSQQHETQTRLLSLV

>C.tovari

MEFDLRYLILILANMVKSISSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRTTTTADLSKSFNI  
LCLASLLLPQMLFWYFFFFFIITLSSCSSWIGNAWASFRQSLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLSLV

>468-2

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRTTTTADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>CaBs4C.1

MEFDLRYFILILANMLKSILSISNNWDPFHIIIGNPIPSFMTFMNRLFFLFIFSFIFSITRMTLHHPNIRIRVSITFSKSFNII  
CLASLLLPQMLFWYFFFIITLSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVISQQHETQTNHLLSLV

>CaBs4C.2

MEFDLRYLILIIIVNILKSISPSENWDPFHIFGNPIPSFITFMNRLAFLFLFSFIFSITRTTLHHPNIRIRISITFSKSFIFSITRT  
TLHHPNIRIRISITFSKSFNII  
CLASLLLPQMLFWYFFFIITLSSCSSWIGNAWASFRQPTLHIFSTIIFPAVSIFINVEFDRNTV  
SQHETQTRLLSLV

>CaBs4C.5

MEFDLRYTILIIIVNILKSIFISNEWDPFHIFVQLTFMAFMNRLFLLLLFSFIFSITSITIHYPMRVHISSTYFSKSFNII  
FLTSFL  
LPQVLFWYIFFITIMTFPSNWNISNFKEWSLYIISIIPGFSIFINAEIDRNNVNNQQPQAQTIHLLLPV

>CpBs4C.2-R

MEFDLRYLILIIIFNILKSISPSENFDPFHIFGNPIPSFITFMNRLAFVFLFSFIFSITRITLHHPNIRIRISITFSKSFNSLCLAS  
LLLPQMLFWYFFFIITLSSCSSWISHTWASVRQPTLHIFSTILPAVSIFINVEFDRNNVSQQHETPTNRLLSLV

>CpBs4C.2-S

MEFDLRYLILIIIVNILKSISPSENFDPFHIFGNPIPSFITFMNRLAFVFLFSFIFSITRITLHHPNIRIRISITFSKSFNII  
CLAS  
LLLPQMLFWYFFFIITLSSCSSWISHTWATVRQPTLHIFSTILPAVSIFINVEFDRNNVSQQHETPTNRLLSLV

>Bs4C-R

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRVRTTTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>Bs4C-S

MEFDLRYFILILANMIKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRTTTTADLSKSFNI  
LCLASLLLPQMLFWYFFFIITLSSCSSWIGNAWASFRGRTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>1480-4

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRVRTTTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>1480-2

MEFDLRYFILILANMIKSILSSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRIRTTTTADLSKSFNI  
LCLASLLLPQMLFWYFFFIITLSSCSSWIGNAWASFRGRTLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>1693-1

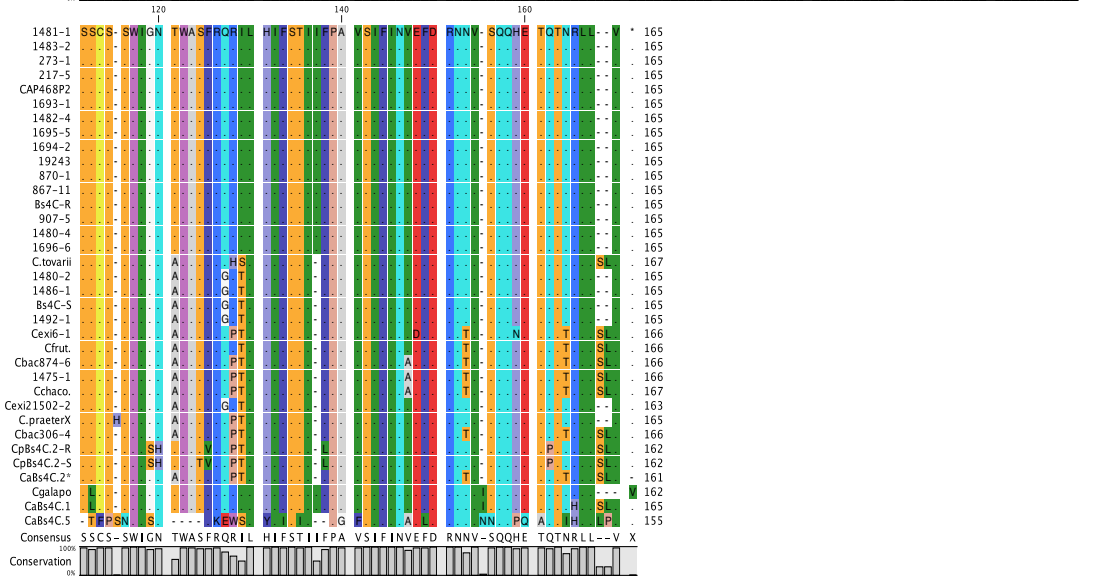
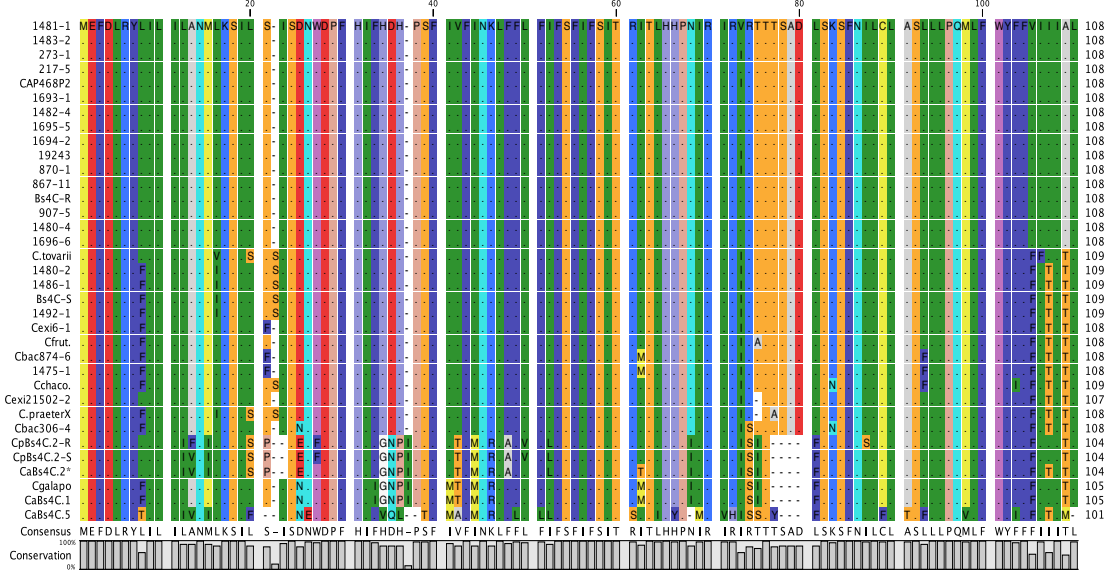
MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRVRTTTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>1695-5

MEFDLRYLILILANMLKSILSISDNWDPFHIFHDHPSFIVFINKLFFLFIFSFIFSITRITLHHPNIRIRVRTTTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIIALSSCSSWIGNTWASFRQRLHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

>1696-6  
MEFDLRYLILANMLKSILSDNWDPFHIFHDHPSFIVFINKLFFLFISFISITRITLHHPNIRIRVRTTTTSADLSKSFNII  
CLASLLLPQMLFWYFFVIIALSSCSSWIGNTWASFRQRILHIFSTIIFPAVSIFINVEFDRNNVSQQHETQTNRLLV

Alignment (A) und Vergleich (B) der Proteinsequenzen der verschiedenen Allele  
A)



B)

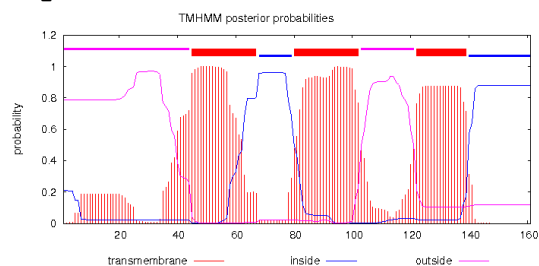
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	
1481-1	1	100,00	99,39	100,00	99,39	100,00	100,00	100,00	100,00	100,00	99,39	99,39	100,00	100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	
1483-2	2	0		99,39	100,00	99,39	100,00	100,00	100,00	100,00	99,39	99,39	100,00	100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	
273-1	3	1	1	99,39	100,00	99,39	99,39	99,39	99,39	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	
217-5	4	0	0	1		99,39	100,00	100,00	100,00	99,39	99,39	100,00	100,00	100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	
CAP68P2	5	1	1	0	1		99,39	99,39	99,39	99,39	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	
1693-1	6	0	0	1	0	1		100,00	100,00	100,00	99,39	100,00	100,00	100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
1482-4	7	0	0	0	1	0	1		100,00	100,00	99,39	100,00	100,00	100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
1695-5	8	0	0	0	1	0	1	0		100,00	100,00	100,00	100,00	100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
1694-2	9	0	0	0	1	0	1	0	0		99,39	100,00	100,00	100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
19243	10	1	1	0	1	0	1	1	1	1		100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	
870-1	11	1	1	0	1	0	1	1	1	1	0		99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	
887-11	12	0	0	1	0	1	0	0	0	0	1	1		100,00	100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
Ba4C-R	13	0	0	1	0	1	0	0	0	0	1	1	0		100,00	100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
907-5	14	0	0	1	0	1	0	0	0	0	1	1	0	0		100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
1480-4	15	0	0	1	0	1	0	0	0	0	1	1	0	0	0		100,00	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
1696-6	16	0	0	1	0	1	0	0	0	0	1	1	0	0	0	0		99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39	99,39
C.brevi	17	13	13	12	13	13	13	13	13	13	12	12	13	13	13	13	13		94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01	94,01
1480-2	18	11	11	10	11	10	11	11	11	11	10	11	11	11	11	11	11	10		100,00	100,00	100,00	93,41	95,21	92,81	92,81	93,41	97,58	96,39	93,41	75,60	75,60	79,17	82,14	81,18	55,62
1486-1	19	11	11	10	11	10	11	11	11	11	10	11	11	11	11	11	11	10	0		100,00	100,00	93,41	95,21	92,81	92,81	93,41	97,58	96,39	93,41	75,60	75,60	79,17	82,14	81,18	55,62
Ba4C-S	20	11	11	10	11	10	11	11	11	11	10	11	11	11	11	11	11	10	0	0		100,00	93,41	95,21	92,81	92,81	93,41	97,58	96,39	93,41	75,60	75,60	79,17	82,14	81,18	55,62
1482-1	21	11	11	10	11	10	11	11	11	11	10	11	11	11	11	11	11	10	0	0	0		93,41	95,21	92,81	92,81	93,41	97,58	96,39	93,41	75,60	75,60	79,17	82,14	81,18	55,62
Cen6-1	22	16	16	15	16	15	16	16	16	16	15	16	16	16	16	16	16	13	11	11	11		96,99	96,99	96,99	95,21	93,37	92,26	96,99	76,65	76,65	82,63	79,29	80,47	55,36	
Chul.	23	13	13	12	13	12	13	13	13	13	12	12	13	13	13	13	13	11	8	8	8	5		96,39	96,39	96,39	95,81	95,78	93,45	96,99	77,25	83,23	81,66	82,84	55,95	
Cha674-6	24	17	17	16	17	16	17	17	17	17	16	16	17	17	17	17	17	14	12	12	12	5	6		100,00	97,01	92,77	91,67	96,76	76,05	82,63	79,98	81,07	56,55		
1475-1	25	17	17	16	17	16	17	17	17	17	16	16	17	17	17	17	17	14	12	12	12	5	6	0		97,01	92,77	91,67	95,78	76,05	82,63	79,98	81,07	56,55		
Chao.	26	18	18	17	18	17	18	18	18	18	17	17	18	18	18	18	18	13	11	11	11	11	8	7	5	5		92,22	92,26	96,41	75,00	80,95	79,24	79,41	56,80	
Cen21502-2	27	9	9	8	9	8	9	9	9	9	8	8	9	9	9	9	9	11	4	4	4	4	11	7	12	12	13		95,15	93,37	77,25	77,25	80,84	82,63	81,66	56,55
C.prieterX	28	15	15	14	15	14	15	15	15	15	14	14	15	15	15	15	15	11	6	6	6	6	13	11	14	14	13	8		92,26	76,92	80,47	81,07	80,12	55,62	
Cha305-4	29	16	16	15	16	15	16	16	16	16	15	16	16	16	16	16	16	13	11	11	11	11	6	5	7	7	6	11	13		77,84	77,84	83,83	81,66	82,84	56,55
CpBa4C2-R	30	40	40	39	40	39	40	40	40	40	39	39	40	40	40	40	40	37	41	41	41	41	39	38	40	40	42	38	39	37		98,15	90,74	81,82	83,03	56,63
CpBa4C2-S	31	40	40	39	40	39	40	40	40	40	39	39	40	40	40	40	40	37	41	41	41	41	39	38	40	40	42	38	39	37	3		91,36	81,82	83,03	57,83
CpBa4C2-S	32	38	37	38	37	38	37	38	38	38	37	37	38	38	38	38	38	33	35	35	35	29	28	29	29	32	32	33	27	15	14		83,64	84,85	59,04	
CaBa4C2'	33	25	25	26	25	26	25	25	25	25	26	26	25	25	25	25	25	34	30	30	30	35	31	34	34	37	29	32	31	30	30	27		97,58	57,49	
Cgalaipo	34	27	27	28	27	28	27	27	27	27	28	28	27	27	27	27	27	32	32	32	32	32	33	29	32	32	35	31	34	29	28	28	25	4		56,68
CaBa4C-1	35	77	77	76	77	76	77	77	77	77	76	76	77	77	77	77	77	76	75	75	75	75	75	74	73	73	73	73	73	73	72	70	68	71	69	



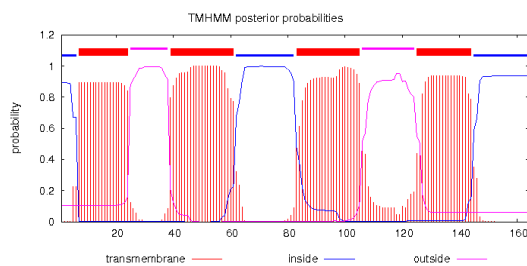


## A 33

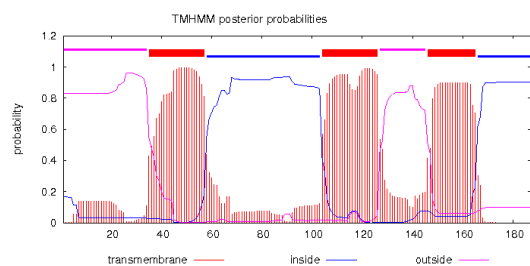
## &gt;CpBs4C.2-R



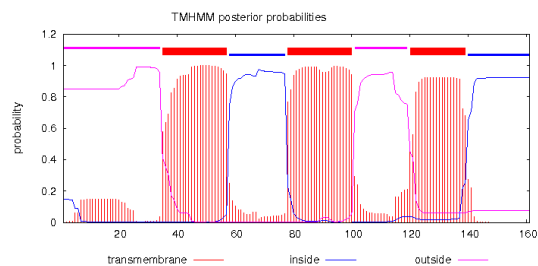
## &gt;Bs4C-S



## &gt;CaBs4C.2



## &gt;CaBs4C.2Del



Transmembran-Helices-Vorhersage nach dem Hidden-Markov-Model für CpBs4C.2-R, CaBs4C.2 und CaBs4C.2Del (Deletionskonstrukt bei welchem die Duplikation entfernt wurde). Die TM-Helices sind als rote senkrechte Striche, und die Polypeptidkette als Linie in Abhängigkeit von ihrer Position, die entweder innerhalb (blau) oder außerhalb (lila) der Membran liegt, dargestellt.

## A 34

Abbildung modifiziert nach Strauß et al., 2012 Tab. S3.

Name	Sequenz 5' -> 3'	Verwendung
12600 fwd1	GCCTGCTACCTGATGTAAGC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd2	CTTACATATTCAGCAGTAGC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd3	CTAGAGATATTCAGTACAGG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd4	GGAGATGTAAGCTGATGAG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd5	CGATATCTACGCTATACAC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd5 rs!	GGATAGGCTACGATGATGAG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd6	GAGTACATCTAAGCTTGCTAG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd7	GATAGAGCATAGAGGATAGGC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd10	CTTCGAGGAACAGCAATGTTA	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd11	CTGATTCATATACATGGCTGGT	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd12	GATGGCAAACTTCCCATG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd13	GACTGATCAACAGCTACAG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd14	GGATCATCAAACTTCTGATCC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd15	CTCAGACATCTAATTAATGAACATC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 fwd16	GACTGCTTATATGATCAACACA	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 rev1	GCTTTACTCAACATCAAGTAGC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 rev2	AGTGTCCGGGCACTGATG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 rev3	GCTACAGTATTTATTTGATGAC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 rev4	GACACATAGTTGATACATGATC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 rev5	GATCCGGTATTAAGTATGATGC	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600 rev6	GAGACAATTCAGGTCAATTCAG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
12600startATG-fwd-E	TTGGTCTCTCAGCATGGAGTTTGAATCTCAGATCTTG	Amplifikation von Bsf4C-R- und Bsf4C-S-KDS für cut-ligation in pENTR-DiBsal-GW
12600+stopP-rev-E	TTGGTCTCACTCTGACGAGCAGCGATTTG	Amplifikation von Bsf4C-R- und Bsf4C-S-KDS für cut-ligation in pENTR-DiBsal-GW
12600-stopP-rev-E / TS	TTGGTCTCACTCTGACGAGCAGCGATTTG	Amplifikation von Bsf4C-R- und Bsf4C-S-KDS (ohne Stoppkodon)
12600startATG-fwd	ATGGAATGATCTCAGATCTG	Amplifikation von Bsf4C-R- und Bsf4C-S-korresp. Sequenzen von genom. DNA
hotbox79in3-fwd	TATAAATAATAGTCTCTCTGTTAAACAATGAACAGCTTGC	Integration des <i>EBE<sub>max</sub>Bsf4C-R</i> in den Bsf3-Promoter
hotbox68in3-fwd	TACAAATAATAGTCTCTCTGTTAAACAATGAACAGCTTGC	Integration des <i>EBE<sub>max</sub>Bsf4C-S</i> in den Bsf3-Promoter
4in308-rev-02	P-GGTGTCGAATTTGGTTAAACC	Integration von EBE in den Bsf3-Promoter
prom79-delbox-fwd	P-AGCCCTGAAACAGAGAAATAATTT	Deletion des <i>EBE<sub>max</sub>Bsf4C-R</i> aus Bsf4C-R-Promoterfragment
prom79-delbox-rev	TAAATATCTCCAGCTCTTAATAATATTTG	Deletion des <i>EBE<sub>max</sub>Bsf4C-R</i> aus Bsf4C-R-Promoterfragment
prom68-delbox-fwd	P-AGCCCTGAGAGCTAGAGATTT	Deletion des <i>EBE<sub>max</sub>Bsf4C-S</i> aus Bsf4C-S-Promoterfragment
prom68-delbox-rev	TAAATATCTCTGCTTAAATAATATTTG	Deletion des <i>EBE<sub>max</sub>Bsf4C-S</i> aus Bsf4C-S-Promoterfragment
12600box1-fwd	TTGGTCTCTCAGCAACAGACAGCTTGC	Amplifikation von Bsf4C-R-Promoterfragmenten
12600-79prom300 fwd	TTGGTCTCTCAGCTCTATCATGCTG TATATTTGTTGTTAG	Amplifikation von Bsf4C-R-Promoterfragmenten
12600box1-rev	TTGGTCTCTCAGCTCTATCATGCTG TATATTTGTTGTTAG	Amplifikation von Bsf4C-R-Promoterfragmenten
semiquantitative RT-PCR-Analysen		
EPrT-F1	AGTCACACACACTGCTAC	sq RT-PCR für <i>EF-1a</i>
EPrT-R1	GGGATGATGACTTAGTGGTC	sq RT-PCR für <i>EF-1a</i>
12600 fwd2	CTTACATATTCAGCAGTAGC	sq RT-PCR für Bsf4C-R und genetische Kartierung
12600 rev1	GCTTTACTCAACATCAAGTAGC	sq RT-PCR für Bsf4C-R und genetische Kartierung
12600 RACE prod fwd2	GATAGATGGAGTTGATCTAG	sq RT-PCR für <i>CpBsf4C.2-R</i> und <i>CpBsf4C.2-S</i>
79-1 specific rev1	GAACTTGGATGGGATGGCC	sq RT-PCR für <i>CpBsf4C.2-R</i> und <i>CpBsf4C.2-S</i>
Oligonukleotide für EMSA- und MST-Analysen		
Bsf4C-1 forward 5' DY682-fwd	CATTTTAATATA AAAAATAGTCTCTCAGCCCTTA	EMSA-Analysen ( <i>Bsf4C-R</i> Oligonukleotid), Cys3-markierter Oligonukleotid für MST-Analysen verwendet
Bsf4C-1 reverse 5' DY682-rev	TTAGGCTGAGAGCTATTTTATATATTA AAATG	EMSA-Analysen ( <i>Bsf4C-S</i> Oligonukleotid), Cys3-markierter Oligonukleotid für MST-Analysen verwendet
Bsf4C-2 forward 5' DY782-fwd	CATTTTAATACCAAAATAGTCTCTCAGCCCTCA	EMSA-Analysen ( <i>Bsf4C-S</i> Oligonukleotid), Cys3-markierter Oligonukleotid für MST-Analysen verwendet
Bsf4C-2 reverse 5' DY782-rev	TTAGGCTGAGAGCTATTTTGTATTAATAATG	EMSA-Analysen ( <i>Bsf4C-S</i> Oligonukleotid), Cys3-markierter Oligonukleotid für MST-Analysen verwendet
Oligonukleotide für RACE-Analysen		
SMARTer II A Oligonukleotide	AAGCAGTGGTATCAACGACGAGTACXXXX	im SMARTER RACE cDNA Kit Clontech enthalten
5'-RACE CDS Primer A	5'-(T)25V N-3'	im SMARTER RACE cDNA Kit Clontech enthalten
Universal Primer A Mix (UPM)	CTAATACGACTCACTATAGGCAAGCTGGTATCAACGACAGT	
Klonierung von blunt end PCR-Produkten in pJET1.2		
pJET seq fwd	CGACTCATATAGGAGAGCGGC	Kol-PCR+Sequenzierung von pJET1.2 nach Insertion (im Kit enthalten)
pJET seq rev	AAGACATCGATTTTCATCGGAC	Kol-PCR+Sequenzierung von pJET1.2 nach Insertion (im Kit enthalten)
pJET Seq F	CTCTTTAACTGCTGGCTGAACACC	Kol-PCR+Sequenzierung von pJET1.2 nach Insertion (neu erstellt von R. Morbitzer)
pJET Seq R	CAACGGTCTGATGAGGTGGTAC	Kol-PCR+Sequenzierung von pJET1.2 nach Insertion (neu erstellt von R. Morbitzer)
Genome Walker-Analysen		
Adaptor Primer 1	GTAATACGACTCACTATAGGCG	Genome Walker-Analysen
Nested Adaptor Primer 2	ACTATAGGCAAGCGGTGGT	Genome Walker-Analysen

M13 fwd	GTAAACGAGCGCCAGT	Kol-PCR-Sequenzierung
M13 rev	GGAAACGCTATGACCATG	Kol-PCR-Sequenzierung
PUC57 fwd	CGATCGGTGGGGCTC	Kol-PCR-Sequenzierung
PUC57 rev	CGGCAGTGAAGCAAC	Kol-PCR-Sequenzierung
Identifizierung von <i>CpBs4C.2-R/S</i>		
12600 fwd12	CGATGCACAACTTTCOAATG	Identifizierung von <i>CpBs4C.2-R/S</i>
12600 fwd12c	CTCTACAATTTCCACGATGAG	Identifizierung von <i>CpBs4C.2-R/S</i>
12600 RACE prod fwd2	GAATAGATGAGTTTATCTCTAG	Identifizierung von <i>CpBs4C.2-R/S</i>
12600-rev1	GCCTTACTCAACATACAAGTCAG	Identifizierung von <i>CpBs4C.2-R/S</i>
Kandidat 630		
630 Prom Fwd1	GAATTGTGTGGTCAATGAT	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 Prom Fwd2	TCAGGTGGTCATCACAGA	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 Prom Fwd2	AGGCTAGGAACAATGGGAT	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 Prom Rev1	TTGAGCATTTTCCCACTTG	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 prom fwd4	GTAGACGAATATCAATAAACAG	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 prom fwd3	CTCCTTACTCCCTGTCTCTTG	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 prom fwd3	GAGTTCGAATCTTGATCAATCC	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
Contig630 rev	CTTGGCAATTCACAAAGCTCAATC	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
Contig630 fwd-RC	GGATTGATCAAGAGATTCGAATC	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 K3 rev	GGAGTTTGGCAATCTTTCG	Amplifikation von Kand. 630-korresp. Sequenzen von genom. DNA
630 prom fwd-E	TTGCTCTCAGCCGCAATAATAATAGTACTGATTGG	Amplifikation von Kand. 630-Promoterfragmenten
630 prom rev-E	TTGGTCTCAGCTTGGCAACAATCGAAGC	Amplifikation von Kand. 630-Promoterfragmenten
Kandidat 15502		
15502 prom fwd 2	GCAAATCAACACCCGCTC	Amplifikation von Kand. 15502-korresp. Sequenzen von genom. DNA
15502 prom fwd 3	GCATCTAATATAGCAATATCAAC	Amplifikation von Kand. 15502-korresp. Sequenzen von genom. DNA
15502 prom rev2	CTCTGGAAATTTGGAGTTTC	Amplifikation von Kand. 15502-korresp. Sequenzen von genom. DNA
Cluster15502 fwd	GCTAGTGAACAGATTTTAAAGC	Amplifikation von Kand. 15502-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster15502 rev	GATCAAGAGCTCATTTGCTGC	Amplifikation von Kand. 15502-korresp. Sequenzen von genom. DNA/sq RT-PCR
Kandidat 8248		
Cluster8248 fwd	CAACTAATCAATGAATCTTTGGAAATG	Amplifikation von Kand. 8248-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster8248 rev	GCACATAGACTTCATTCATTTCAAG	Amplifikation von Kand. 8248-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster 8248end fwd	GTGTAAATAGAAGTAAGAGCAG	Amplifikation von Kand. 8248-korresp. Sequenzen von genom. DNA
Cluster 8248end rev	GTGCTTCATCTTATATATCAAC	Amplifikation von Kand. 8248-korresp. Sequenzen von genom. DNA
8248in3	TATTAATATTAGTCCCTCCCTGGTTAAACAATGACAGCTTGG	Integration des <i>BBF<sub>Wass</sub></i> 8248 in den <i>Bs3</i> -Promoter
Cand2/3		
Bs4 Cand2 fwd	ACGAGTTGGACGACGAAGT	Amplifikation von Kand. 2/3-korresp. Sequenzen von genom. DNA/sq RT-PCR
Bs4 Cand2 rev	CGAGCTATCACACCTTAATGG	Amplifikation von Kand. 2/3-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand4		
Cand4-fwd new	GAACCCAGATGAGGTTGAA	Amplifikation von Kand. 4-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand4-rev	TGGACAGACTTCCAATTTT	Amplifikation von Kand. 4-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand4fwd new2	GTAGTAATTGGGAATAGTAGG	Amplifikation von Kand. 4-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand4-rev new	GGGAGCAGTATAGATGTTGC	Amplifikation von Kand. 4-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand5		
Cand5-fwd	AGAGCAACACCCACAAC	Amplifikation von Kand. 5-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand5-rev	TGGAAATGGCTAARAGATGC	Amplifikation von Kand. 5-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand6		
Cand6-fwd	GTGAATCTGATGGGATGG	Amplifikation von Kand. 6-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand6-rev	ATATTGGCATGGAGTGGAA	Amplifikation von Kand. 6-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand6-fwd new	GCCAATTAACCTTGGATGGAGTTG	Amplifikation von Kand. 6-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand6-rev new	GGTGATAAGGAGTGGTAGC	Amplifikation von Kand. 6-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand6-fwd new2	GAACCTAGAAACACACTAGG	Amplifikation von Kand. 6-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand6-rev new2	GGAAGAAAGCTAAGCACTGAG	Amplifikation von Kand. 6-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand7		
Cand7-fwd	GCTACAATAGTCAAGTTACAAGATAGCG	Amplifikation von Kand. 7-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand7-rev	CCCAAGCAAAAATCTTCAACAGCC	Amplifikation von Kand. 7-korresp. Sequenzen von genom. DNA/sq RT-PCR

Cand9		GTGTCATTTAGAGATCATAC	Amplifikation von Kand. 9-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand9-fwd		CCGCTACTATCCGAGTCCA	Amplifikation von Kand. 9-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand9-rev		GGATCAATTAATACTACTATCAATAACAAG	Amplifikation von Kand. 9-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand9-fwd new		CCATCATGGCAAGTTTAAAGCATGG	Amplifikation von Kand. 9-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand9-rev new		CAGTCGCTCTTTGGCTATATGG	Amplifikation von Kand. 9-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand9-rev new2		GAGTCATCAGCTTCACATCATG	Amplifikation von Kand. 9-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand9-rev new3			
Cand10/11			
Cand10/11fwd		CTGAACATGTAAGATATATTATTAC	Amplifikation von Kand. 11-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand10/11rev		GGTTTGTGGATGGGCTTT	Amplifikation von Kand. 11-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand15			
Cand15-fwd		CCACAGGAGATGGGTGTTT	Amplifikation von Kand. 15-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand15-rev		CAGATGGAGGATGTGGATG	Amplifikation von Kand. 15-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand15-fwd new		GCTTACCACCAAGTATGACGAGAGTAC	Amplifikation von Kand. 15-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cand15-rev new		GAATGACAGAAATGATTTCGCCG	Amplifikation von Kand. 15-korresp. Sequenzen von genom. DNA/sq RT-PCR
Contig1919			
contig1919 fwd		CAGCGCTTCAACATTATTATATAGC	Amplifikation von Kand. 11-korresp. Sequenzen von genom. DNA/sq RT-PCR
contig1919 rev		CACAAGCTCTTCATGAGTTAAAG	Amplifikation von Kand. 11-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster8984			
Cluster8984 fwd		CACAAACCAATTCAATTCAAGAGAAAAAG	Amplifikation von Kand. 8984-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster8984 rev		GATCAGTAAGTGGTACGTCCACATG	Amplifikation von Kand. 8984-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster11620			
Cluster11620 fwd		CAATPACAGTAACAACCATAGCACCC	Amplifikation von Kand. 11620-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster11620 rev		GAATATGCTCGGTATGCGTTG	Amplifikation von Kand. 11620-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster699			
Cluster699 fwd		GGCAACAGCATTTGAAAAGTAAACTTG	Amplifikation von Kand. 699-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster699 rev		GCACAAGATTCATGTGATTGACG	Amplifikation von Kand. 699-korresp. Sequenzen von genom. DNA/sq RT-PCR
cluster 9054			
Cluster 9054 fwd		CACAAAAATCAACATCATACAGTGC	Amplifikation von Kand. 9054-korresp. Sequenzen von genom. DNA/sq RT-PCR
Cluster 9054 rev		CGCTTCGGGTTTTATTCTTTGGTT	Amplifikation von Kand. 9054-korresp. Sequenzen von genom. DNA/sq RT-PCR
Amplifikation der <i>C. annuum</i> -Homologen			
CaBs4C.1			
011445prom fwd		GGGGCTACGAAAACAATTATCC	Amplifikation von CaBs4C.1-Promoterfragmenten
011445stopp-rev-E		TTTGTGTCACCTTCCTATGACGAGAGAGACAGTG	Amplifikation von CaBs4C.1-korresp. Sequenzen von genom. DNA
011445-stopp-rev-E		TTTTGCTCTCACCCTGACGAGAGAGAGTGATTTG	Amplifikation von CaBs4C.1-korresp. Sequenzen von genom. DNA
011445ATG-fwd-E		TTTTGTGCTCTCACCATGAGTTGATCTAAGATATTC	Amplifikation von CaBs4C.1-korresp. Sequenzen von genom. DNA
011445prom fwd-E		TTTTGTGCTCTCACCAGGCTACGAAACAATTATTC	Amplifikation von CaBs4C.1-Promoterfragmenten
011445prom rev-E		TTTTGTGCTCACTTCATATATCTGTTGAATCTTTCC	Amplifikation von CaBs4C.1-Promoterfragmenten
C.a.011445boxin3		TACAAAAATAGCCCTCCCTGGTTAAACATGACACGTTGC	Integration des <i>EBE<sub>Yms1</sub></i> -CaBs4C.1 in den <i>Bs3</i> -Promoter
CaBs4C.2			
082548prom fwd		AACTACAGTACATCATATGGGCTGG	Amplifikation von CaBs4C.2-Promoterfragmenten
082548 3'UTR rev2		GTACACAGACCCATAAAATTTTAC	Amplifikation von CaBs4C.2-Promoterfragmenten
082548-3'UTR rev1		CACATATATGAAATPACAAAGCTTG	Amplifikation von CaBs4C.2-Promoterfragmenten
082548-stopp-rev-E		TTTTGTGCTCACTTCCTACGACGAGACGAGCGCAG	Amplifikation von CaBs4C.2-korresp. Sequenzen von genom. DNA
082548-stopp-rev-E		TTTTGTGCTCACTTCCTACGACGAGACGAGCGCAG	Amplifikation von CaBs4C.2-korresp. Sequenzen von genom. DNA
082548-stopp-rev-E		TTTTGTGCTCACTTCCTACGACGAGACGAGCGCAG	Amplifikation von CaBs4C.2-korresp. Sequenzen von genom. DNA
082548-stopp-rev-E		TTTTGTGCTCACTTCCTACGACGAGACGAGCGCAG	Amplifikation von CaBs4C.2-korresp. Sequenzen von genom. DNA

## DANKSAGUNG

Bei Thomas möchte ich mich für die Aufnahme in die Arbeitsgruppe danken. Aus dem „ich möchte mal ein bißchen reinschnuppern was ihr so macht“ sind dann doch ein paar Jahre geworden. Danke für das spannende Thema, die Betreuung und die Unterstützung vor allem als der 12 Stunden-Ansatz doch ein bißchen zeitig war.

Bei Prof. Dr. Martin Parniske möchte ich mich für die Erstellung des Gutachtens bedanken sowie die freundliche und offene Aufnahme in die Genetik nach unserem Umzug von Halle nach München.

Insbesondere bei Familie Strauß (A<sup>3</sup>), möchte ich mich bedanken, dass ihr mich auch in den nicht so leichten Phasen dieser Arbeit unterstützt haben und immer ein offenes Ohr hatten. Annett, danke für deine hilfreichen Kommentare, das Korrekturlesen, deine Geduld, deine Tipps und für soviel mehr. Ich habe eine Menge von Dir lernen dürfen und das schätze ich sehr.

Allen ehemaligen und derzeitigen Mitgliedern der Arbeitsgruppe Lahaye (Halle und München) möchte ich mich für die angenehme Arbeitsatmosphäre, die Kuchen- und Obstpausen, die lustigen Labortage, die kleinen und großen Hilfen im Labor sowie für die Meetings und Diskussionsrunden bedanken. Danke an Sebastian Schornack von dem ich viel lernen durfte als Hiwi, vor allem, dass man den Phänotyp nur wollen muss. Danke an Patrick, der seine Chicks immer im Griff hatte. Danke an Robert für die vielen Gespräche während der Gewächshausnachmittage und jetzt isses doch schön geworden. Und viele andere liebe Leute aus dem Labor, die die Zeit auf Arbeit positiv gestaltet haben: Annett, Janett, TomS (1-2-3-4), Christina W., Orlando, Sabine R., Christina K., Beate, JensH, Niklas, Jessy und Corinna. Den Ottis aus dem Sunshinelab (coole Truppe), der JessyF. und SimoneB. Ein Dankeschön geht auch an all die guten Seelen, wie dem one and only Karl Heinz (es war mir eine Freude und Ehre), an Sequenzier-Gisela (und AndreasB), die Spülfrauen, die Gärtnern, die Werkstatt und alle anderen technischen Angestellten ohne deren fleißiges Tun vieles beschwerlicher wäre.

Ein Dank geht auch an alle meine Freunde, die mir die Zeit außerhalb des Laboralltags abwechslungsreich gestaltet haben und die immer da sind wenn man sie braucht.

Bei meiner Familie möchte ich mich für die immerwährende Liebe und Unterstützung bedanken und Euch sagen wieviel mir das bedeutet-DANKESCHÖN! Ohne Euch wäre das nicht möglich geworden.